

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Publications from USDA-ARS / UNL Faculty

U.S. Department of Agriculture: Agricultural
Research Service, Lincoln, Nebraska

2011

26th Hohenheim Consensus Conference, September 11, 2010 Scientific substantiation of health claims: Evidence-based nutrition

Hans Konrad Biesalski
University of Hohenheim, biesal@uni-hohenheim.de

Peter J. Aggett
Lancaster University

Robert Anton
University of Strasbourg

Paul S. Bernstein
University of Utah School of Medicine

Jeffrey Blumberg
Tufts University

See next page for additional authors
Follow this and additional works at: <https://digitalcommons.unl.edu/usdaarsfacpub>

 Part of the [Agricultural Science Commons](#)

Biesalski, Hans Konrad; Aggett, Peter J.; Anton, Robert; Bernstein, Paul S.; Blumberg, Jeffrey; Heaney, Robert P.; Henry, Jeya; Nolan, John M.; Richardson, David P.; van Ommen, Ben; Witkamp, Renger F.; Rijkers, Ger T.; and Zöllner, Iris, "26th Hohenheim Consensus Conference, September 11, 2010 Scientific substantiation of health claims: Evidence-based nutrition" (2011). *Publications from USDA-ARS / UNL Faculty*. 836.
<https://digitalcommons.unl.edu/usdaarsfacpub/836>

This Article is brought to you for free and open access by the U.S. Department of Agriculture: Agricultural Research Service, Lincoln, Nebraska at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Publications from USDA-ARS / UNL Faculty by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors

Hans Konrad Biesalski, Peter J. Aggett, Robert Anton, Paul S. Bernstein, Jeffrey Blumberg, Robert P. Heaney, Jeya Henry, John M. Nolan, David P. Richardson, Ben van Ommen, Renger F. Witkamp, Ger T. Rijkers, and Iris Zöllner



Special article

26th Hohenheim Consensus Conference, September 11, 2010 Scientific substantiation of health claims: Evidence-based nutrition

Hans Konrad Biesalski Ph.D.^{a,*}, Peter J. Aggett Ph.D.^b, Robert Anton Ph.D.^c,
Paul S. Bernstein M.D., Ph.D.^d, Jeffrey Blumberg Ph.D.^e, Robert P. Heaney Ph.D.^f,
Jeya Henry Ph.D.^g, John M. Nolan Ph.D.^h, David P. Richardson Ph.D.ⁱ, Ben van Ommen Ph.D.^j,
Renger F. Witkamp Ph.D.^k, Ger T. Rijkers Ph.D.^l, Iris Zöllner Ph.D.^m

^a Institute of Biological Chemistry and Nutrition, University of Hohenheim, Stuttgart, Germany

^b Lancaster University, Parbold, United Kingdom

^c University of Strasbourg, Strasbourg, France

^d Department of Ophthalmology and Visual Sciences, Moran Eye Center, University of Utah School of Medicine, Salt Lake City, Utah, USA

^e Antioxidants Research Laboratory, Jean Mayer USDA Human Nutrition Research Center on Aging, Friedman School of Nutrition Science and Policy, Tufts University, Boston, Massachusetts, USA

^f Creighton University, Omaha, Nebraska, USA

^g Functional Food Centre, Oxford Brookes University, Oxford, United Kingdom

^h Macular Pigment Research Group (MPRG), Department of Chemical and Life Sciences, Waterford Institute of Technology, and Institute of Vision Research, Whitfield Clinic, Waterford, Ireland

ⁱ DPR Nutrition Limited, Croydon, Surrey, United Kingdom

^j TNO, Zeist, The Netherlands

^k Department of Nutrition and Pharmacology, Division of Human Nutrition, Wageningen University, Wageningen, The Netherlands

^l Department of Surgery, University Medical Center, Utrecht, and Laboratory of Medical Microbiology and Immunology, St. Antonius Hospital, Nieuwegein, The Netherlands

^m Baden-Wuerttemberg State Health Office, Epidemiology, Stuttgart, Germany

ARTICLE INFO

Article history:

Received 5 April 2011

Accepted 6 April 2011

Keywords:

Nutrition
Clinical trials
Health claim
Micronutrients
Regulation

ABSTRACT

Objective: The objective was to define the term evidence based nutrition on the basis of expert discussions and scientific evidence.

Methods and procedures: The method used is the established Hohenheim Consensus Conference. The term “Hohenheim Consensus Conference” defines conferences dealing with nutrition-related topics. The major aim of the conference is to review the state of the art of a given topic with experts from different areas (basic science, clinicians, epidemiologists, etc.). Based on eight to 12 questions, the experts discuss short answers and try to come to a consensus. A scientifically based text is formulated that justifies the consensus answer. To discuss the requirements for the scientific substantiation of claims, the 26th Hohenheim Consensus Conference gathered the views of many academic experts in the field of nutritional research and asked these experts to address the various aspects of a claims substantiation process and the possibilities and limitations of the different approaches.

Results: The experts spent a day presenting and discussing their views and arrived at several consensus statements that can serve as guidance for bodies performing claims assessments in the framework of regulatory systems.

Conclusion: The 26th Hohenheim Consensus Conference addresses some general aspects and describes the current scientific status from the point of view of six case studies to illustrate specific areas of scientific interest: carotenoids and vitamin A in relation to age-related macular degeneration, the quality of carbohydrates (as expressed by the glycemic index) in relation to health and well-being, probiotics in relation to intestinal and immune functions, micronutrient intake and maintenance of normal body functions, and food components with antioxidative properties and health benefits.

© 2011 Published by Elsevier Inc.

This work was financially supported by the European Responsible Nutrition Alliance.

* Corresponding author. Tel.: +49-0-711-459-24112; fax: 49-0-711-459-23822.

E-mail address: biesal@uni-hohenheim.de (H. K. Biesalski).

Introduction

An important objective of global regulatory legislation is to ensure that health claims on foods and food constituents can be properly justified and scientifically substantiated. Consumers should be able to make choices based on clear and accurate information and have confidence in the scientific and regulatory processes used to support health claims. Any framework must not only protect consumers from false and misleading claims but also satisfy the needs of industry to promote and innovate, to stimulate multidisciplinary research, and to reinvigorate efforts to process and preserve raw materials from agriculture, horticulture, fisheries, and aquaculture into a diverse range of foods and food supplements.

Global regulatory frameworks currently reflect the role of food and food constituents in maintaining and promoting human health and in decreasing the risk of disease. The purpose of this consensus conference was to consider recent developments and initiatives on the scientific substantiation of health claims and to pay particular attention to the following items.

- The assessment of the totality of the available data and the need for the development of a scientific framework for weighing the strength, consistency, biological plausibility, and coherence of the evidence.
- A critical examination of the application of the model for evidence-based medicine in nutrition science; evidence-based medicine is designed to evaluate the effects of drugs and not the unique properties, complex effects, and interactions of nutrients and bioactive substances.
- The need to address and define evidence-based nutrition, to embrace state-of-the-art nutrition science, and to stimulate future academic research.
- The identification and validation of relevant biomarkers that can predict potential benefits relating to maintenance or improvement of a function and those associated with decreased risk of disease.
- The determination of the strength of recommendations to reflect the available evidence and consensus among experts in a particular field to communicate truthful nutritional and health messages to the public.

The final decision on whether or not a claim should be allowed rests with the regulators and policy managers, taking into account the scientific assessments of the strength and coherence of the evidence and other societal and health policy considerations including the broader purposes of food labeling and claims to help and improve the public understanding of nutrition and to support strategies for public health promotion and improvement.

To discuss the requirements for the scientific substantiation of claims, the 26th Hohenheim Consensus Conference gathered the views of many academic experts in the field of nutritional research and asked these experts to address the various aspects of a claims substantiation process and the possibilities and limitations of the different approaches. The experts spent a day presenting and discussing their views and arrived at some consensus statements that can serve as guidance for bodies performing claims assessments in the framework of regulatory systems. This report addresses several general aspects and describes the current scientific status from the point of view of five case studies to illustrate specific relevant perspectives:

- Carotenoids and Vitamin A in relation to age-related macular degeneration (AMD).

- The quality of carbohydrates (as expressed by the glycemic index [GI]) in relation to health and well-being.
- Probiotics in relation to intestinal and immune functions.
- Micronutrient intake and maintenance of normal body functions.
- Food components with antioxidative properties and health benefits.

I. Assessing the totality of the available evidence for diet–health relations

This section discusses methodologies for assessing the validity of diet–health relations, with a particular focus on how to consider the totality of the available data, weigh the evidence, and address the strength, consistency, and biological plausibility of a particular diet and health relation. It also deals with requirements and processes to develop research methodologies to demonstrate beneficial effects in a healthy population.

Consensus statements

There is a need for a scientific framework for the assessment of health claims describing the methodology on how to consider the totality of the evidence and weighing the strength, consistency, and biological plausibility of the available evidence.

The Process for the Assessment of Scientific Support for Claims on Foods (PASSCLAIM) described criteria and standards by which the quality and relevance of the scientific evidence should be presented for evaluation and thus the extent to which claims based on them can be said to be scientifically valid. PASSCLAIM provided a scientific framework to facilitate the assessment of scientific support for claims on foods, but the project did not specifically address how the totality of the evidence should be weighed.

A key question is how best existing and new methods can be used to determine transparently the weight of evidence in such a way that the totality of the available data can be evaluated. There are many methods for weighing the evidence (Appendix) that can be used to show the extent to which a cause-and-effect relation is determined, taking into account the nature, quantity, and quality of the different sources of evidence. The scientific framework and criteria for substantiation of claims should be applied intelligently and sensitively to existing and potential claims on a case-by-case basis with respect to gaps in existing knowledge and to the development of new knowledge.

There is a need to define “generally accepted scientific evidence.” There is a need for broad more consensus-based agreements on the scientific substantiation of claims, i.e., for bringing together the experts in the field and let them reach agreement on the totality of the available data and weight of the evidence based on their accepted opinion that is generally held to be correct.

There is a need to translate broad fields of health (e.g., complex physiologic functions such as immune and gastrointestinal functions) into a framework of (bio-)markers and other evidence that would be required to sufficiently substantiate the effect for the maintenance of the function.

The terminologies of “grading of evidence,” which is widely used for the development of nutritional recommendations and which may result in differentiation of claims wordings dependent on the strength of the supporting evidence, constitute a valuable concept that needs to be further explored and applied.

1. What approaches exist and are used to consider the totality of the available data, weigh the evidence, and address the strength,

consistency, and biological plausibility of the relation between diet and health?

Background

The public's increasing awareness that nutrition and dietary components contribute significantly to personal well-being and health, and to public health overall, opened up opportunities and a need to make claims about the benefits of particular foods or of foods components. In consequence, there was an increase in claims about the relations between the consumption of particular food categories, foods, or food constituents and a specific benefit to health. These claims were usually designed to promote a particular product in the face of market competition. The quality of the evidence underpinning such claims was variable, although some claims derived from generally accepted knowledge, which may not be as sound as nutritional science accepts. Others were based on selective and non-robust extrapolations from observational studies, notions that larger intakes of components must endow correspondingly more benefit, folkloric associations, and perhaps more from aspiration rather than empiricism [1,2]. Thus a range of products appeared spanning a spectrum of traditional food products that had been fortified or enriched with specific components to completely novel products. This was the start of the era of what became called "functional foods."

These developments coincided with national and international reviews of reference intakes for nutrients. These reviews demonstrated the limitations of the information from which reference intakes were induced, and they highlighted a need to strengthen the science base. There emerged a movement, evidence-based nutrition, akin, but not necessarily analogous to evidence-based pharmacology and evidence-based medicine. Although evidence-based nutrition and evidence-based medicine are not fully analogous, both benefit from a core objectivity and criticality, and it was appreciated that more hypothesis-based approaches were needed in nutritional science. This particularly applied within the global food economy, which was becoming imbalanced by the use of unregulated and unsubstantiated claims [2]. In consequence, the European Union funded an International Life Sciences Institute (ILSI) that hosted a Concerted Action on Functional Science in Europe (FUFOS), the purpose of which was to characterize the evidence against which a functional effect could be judged [1]. Thus FUFOS met the needs of those developing new products with a focus on very specific health benefits and strong claims and of those who wished to see more rigorous evidential approaches adopted within nutritional science (see below).

Causality and claims

Causality and a claim are two different things. Causality clearly is an integral component of justifying a claim, but causality and claims need to be differentiated from each other. FUFOS and PASSCLAIM deal with the evidence for causality, i.e., the demonstration of the extent to which a cause-and-effect relation is established. The presented evidence of causality is subjected to a scientific appraisal by an independent and competent agency. This independent assessment should then, as was stated in the INTRODUCTION, be reviewed in a broader context by policy makers and regulators in the regulatory authority that allows claims.

Thus the establishment of a claim comprises two elements. The first is the assessment of the evidence to support the extent

to which a causal relation has been demonstrated between a food or food constituent at a specific exposure and a health outcome of interest, and the second is the use of this information in the context of other appropriate considerations to determine whether or not a claim might be justified and what the precise character of that claim might be, i.e., whether it should be qualified in any way.

Thus the assessment of the extent of causality involves a consideration of all the available information against criteria that depend on the causal relation being justified. Although it may not be essential to know the underlying mechanism for this relation, having one and a related hypothesis would appreciably help the selection, quality assurance, validation, compilation, and presentation of the evidence.

Demonstrating causality: quality of the evidence

The essence of good evidence is that it uses objective and valid information suitable for the intended purpose. Enough evidence is the amount that establishes the clearest assessment of the probability of an association between a specific exposure or event and a specific outcome. When assessing the validity of a claim, the reviewing bodies should have access to and consider on their scientific merit all relevant data including those studies that may provide conflicting evidence.

The quality of the markers and outcomes used in substantiating a causal relation is crucial and their quality characteristics are common to all good scientific experimental practice. FUFOS explained these in the context of nutritional science and provided a strategy for their use in nutritional interventional and observational studies, and it assumed that it should be possible to envisage and possibly show the chain of events from eating the component of interest to the final outcome [1]. Integral to this was the use of markers for 1) consumption or external exposure, 2) internal exposure or body burden, 3) intermediate metabolism and the production of active metabolites with, if possible, a quantitative assessment thereof, these markers would represent target function and biological responses, 4) intermediate endpoints, and, if possible, 5) an endpoint directly identified with a decreased risk of disease (Fig. 1).

Ideally evidence, based on a series of markers, could be presented to justify a causal chain of events from early exposure to the desired outcome measurement. However, as Figure 1 illustrates, each event could lead to other events that are not related to the outcome of interest. Furthermore, there may be limited evidence of links between the various "steps," and each hypothetical step may have many other components, so a simplified figure was developed to represent and summarize the key outcome of FUFOS in terms of the types of markers [1] (Fig. 2).

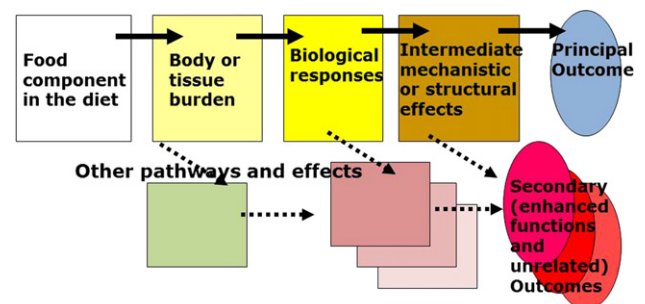


Fig. 1. The chain of markers: consumption to outcome. Principle of a pathophysiologic pathway.

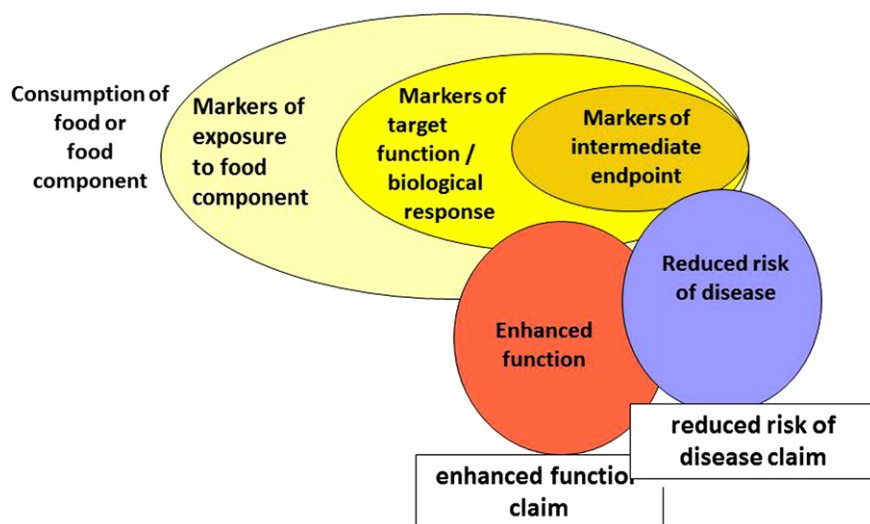


Fig. 2. Concerted Action on Functional Science in Europe: from scientific evidence based on markers for functional foods to types of claims relevant to them [1].

In addition, FUFOS emphasized the principle that markers representing an event directly (i.e., causally) involved in the beneficial or potentially pathogenic process could be regarded as *factors*, whereas those representing correlated but not identifiably causative events should be considered *indicators*.

The more temporally or mechanistically remote a marker is from the target endpoint, the less specific and more attenuated and subject to confounding variables it becomes. Conversely, the closer the marker is to the endpoint in question, the more likely it is to be specific and quantitatively related to the outcome and as such to be a factor rather than an indicator. Clearly the elucidation of the mechanisms leading to health outcomes would refine the identification of markers and their use in experimental and observational studies.

The FUFOS made the following additional points about markers [1]. Overall, markers should be feasible, valid, reproducible, sensitive, and specific. These qualities are applicable to markers whatever their nature (see below). In an experimental study markers should represent relatively immediate outcomes, which can be used to assess interventions in a reasonable time scale; they could, therefore, wherever possible, replace later and more remote outcomes as have been used in some epidemiologic studies. They need to be subjected to the accepted standard procedures for quality assurance. They must also be validated in the context in which they are being used, i.e., to measure dietary intake, internal burden, or intermediate or final outcomes. Markers are also more credible if they have undergone studies to establish their sensitivity (i.e., the frequency of a negative test when the process is present) and their specificity (i.e., the frequency of a positive test when the process is absent), and they must be shown to be reproducible in different centers.

Understandably markers would need to be measured or derived from easily accessible material or obtainable using methodologies that must be ethical and minimally invasive.

The FUFOS consensus eschewed the term *biomarker*, which has biochemical connotations. It preferred *markers* because this acknowledged the broad range of available and applicable markers. Thus, apart from biochemical markers, study outcomes can quite feasibly be derived from, among others, behavioral or psychometric outcomes, physiologic performance, adaptive phenomena, and metabolic clearance studies; indeed, dynamic

responses might be as useful as, or more useful than, static measurements. In fact, any marker that can be quality assured and validated should be appropriate for establishing causality.

Often, it is likely that a “battery” of markers might be needed to address the perspectives of evidence appraisal outlined below for multiple and variable sources of data [3]. In fact, given the quality of data available for health claims, these considerations about markers should inform “new human intervention studies using appropriate markers to generate readily interpretable, valid and reliable data” [1].

One can accept that the FUFOS strategy of using markers from a sequential path of events leading to the hypothesized benefit or decreased risk is an ideal. Nonetheless, it epitomized the need for a strategic approach to the acquisition and analysis of data, and it emphasized the need for and the validation and quality assurance of data and its overall structured integration. Thus these principles would be relevant to the design of studies to obtain new data and to the evaluation of existing data. FUFOS provides a way to assess the quality and relevance of the markers reported in studies and to evaluate the individual reports and sources of information before they are included in a pool of total evidence for collective evaluation and presentation. Interpretation and categorization of information according to the FUFOS strategy may also enable that information to demonstrate its compatibility with Hill’s principles (see below).

All available information needs to describe all the identifiable uncertainties and variabilities affecting its interpretation (e.g., Fig. 3). Clearly these factors influence confidence in the subsequent assessment of causality and, in the context of regulatory affairs, might be foreseen to influence the qualification of a claim and any assessment of safe upper limits of exposure to the food or food component.

Although this section has been written with human studies and applications in mind, it has not considered which data could be derived from animal or other model systems and which would be needed for human studies. The FUFOS and the subsequent concerted action, PASSCLAIM, considered this and appreciated that markers, mechanisms, and plausibility could be explored in systems in vitro and in vivo models and that key data would be generated by appropriate human studies [1].

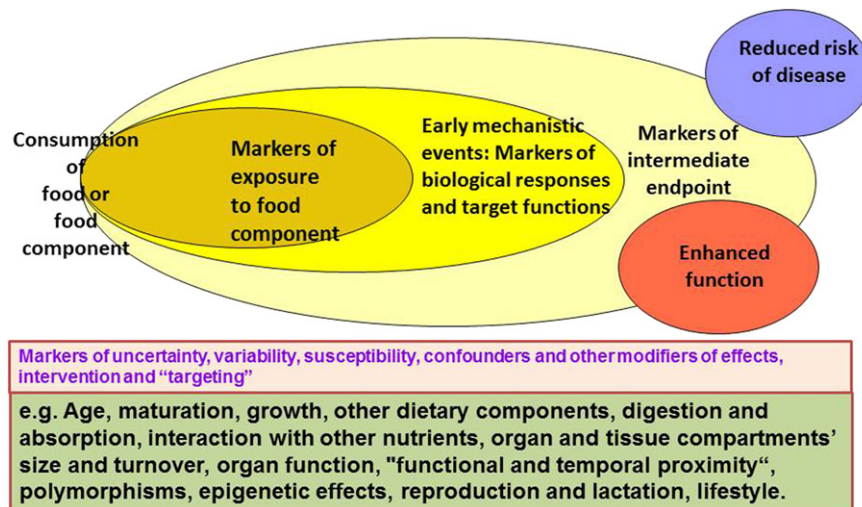


Fig. 3. Modifiers of effect and confounders of the evidence base and markers to explore causality (developed from the Concerted Action on Functional Science in Europe schema [1]).

Demonstrating causality: types of evidence

In Europe, a subsequent concerted action, the PASSCLAIM, also hosted by ILSI Europe, produced a generic tool for assessing the scientific support for health claims of foods and established criteria for markers that could be used to explore the links between diet and health. The latter PASSCLAIM criteria endorsed the FUFOS proposals on markers and the strategy for their use. It also provided guidance on the overall portfolio of evidence submitted to support a health claim [4]. Essentially, PASSCLAIM stated that the portfolio should provide information on the characteristics of the food component or food under study and that the portfolio should contain evidence from human studies, preferably prospective intervention studies, in which 1) the study groups are representative of the target group and appropriately "controlled"; 2) there is an adequate duration of exposure and follow-through demonstrate the intended effect; 3) the study group's background and exposure and other relevant aspects of lifestyle are well characterized; 4) the amount of food or food component eaten is consistent with its intended pattern of use and the influence of the food and dietary matrix is known; 5) the participants' compliance is monitored; and 6) the statistical power is appropriate to test the hypothesized causality. The PASSCLAIM criteria re-emphasized the quality criteria of the markers used and the ideal need for dose-related biologically appropriate changes in their values.

The PASSCLAIM emphasized that claims on foods should be scientifically substantiated by taking into account the totality of the available data and by weighing the evidence. It did not provide specific guidelines as to how the data should be weighed. Its advice constituted a guidance template to inform the evaluative and regulatory processes and to enable the assessment of the appropriateness of the received data for assessment. The evidence supporting the extent to which a causal relation can be inferred comprises five categories [5]. These are 1) experiments that probably can provide the strongest idea of mechanisms with an immediacy between exposure and outcome; 2) randomized controlled trials (RCTs); 3) regression continuity designs; 4) natural experiments (random non-manipulative biomedical occurrences in populations that can be quantitatively and systematically analyzed); and 5) non-experimental studies including cohort studies, case-control

studies, and ecological designs. All of these, with or without corroborative information from animal studies can be used to substantiate the extent to which causality can be demonstrated [5]. PASSCLAIM deliberately did not rank these types of evidence, in part because it was appreciated that each portfolio of evidence would need to be considered in its own right and that this process would probably involve different approaches to the interpretation of the evidence and the uncertainties therein. Note, however, the importance given in the prioritization to studies that indicate mechanistic associations.

Thus PASSCLAIM set what could be regarded as a gold standard for the portfolio of evidence, but by not giving precedence to any particular form of information. It left opportunities for an informed and intelligent compositing and appraisal of the data free from any distortions that might arise from giving greater weight to some lesser-value information because it came from a more highly ranked type of study. Even so, experience suggests that current assessments depend very much on evidence from RCTs. This was not PASSCLAIM's intention. Unfortunately, PASSCLAIM made no comment on how the overall totality of the evidence should be weighed and interpreted. The possibilities and limitations of the use of RCTs in the assessment of the scientific evidence for diet–health relations are further discussed in section III.

Demonstrating causality: totality of the evidence

Philosophically, nothing can be absolutely proved, but a qualified and informed inference about the extent of causality may be induced from appropriately constructed and collated evidence. For individual studies and for quantitative data, probability values infer a degree of causality. However, association does not necessarily establish causality. Sir Austin Bradford Hill noted that tests of significance (probability) inform on the size of effects, but that as such they do not necessarily prove causality [3]. He proposed nine viewpoints from which evidence of associations should be regarded and explored before deciding that any association reflects causality.

1. Strength of the association: A close association is more likely to indicate a causal relation but a slight association does not exclude this possibility.

2. Consistency: Consistent and reproducible associations particular from different centers increase the likelihood of a causative effect.
3. Specificity: Causative relations are more probable if there is strong association between a specific agent and a particular outcome.
4. Temporality: The outcome of interest must occur after the specific exposure.
5. Biological gradient: There needs to be a clear dose–response relation between consumption (external exposure) and/or body burden (internal exposure) to the food component of interest and the outcome. The relation may be direct or inverse. It should be remembered that this relation is subject to uncertainties arising from, among others, absorptive efficiency, bioavailability, homeostasis, and metabolic adaptation.
6. Plausibility: It would be helpful if any observed association is biologically plausible.
7. Coherence: Interpreting data as evidence of causation needs to be considered in the context of other epidemiologic and laboratory findings and knowledge. If it fits, this increases the likelihood of an effect. However, lack of a fit and absence of other information does not negate causation.
8. Experiment: Hill's viewpoints addressed observational data, often from retrospective studies and non-interventional data; thus the availability of well-designed experimental studies in humans and animal models would improve the coherence and plausibility within a portfolio of data.
9. Analogy: The effect of similar factors in other foods and the diet may be considered as part of the overall scenario of data.

Hill wrote that “none of my nine viewpoints can bring indisputable evidence for the cause and effect hypothesis and none can be required as a *sine qua non*. What they can do, with greater or less strength is to help us make up our minds on the fundamental question—is there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect?” [3]. The nine viewpoints are not a checklist; indeed, one of the reasons for setting the viewpoints was to provide an approach to the interpretation of portfolios of data and to provide a way to compensate for data gaps.

It can be easily appreciated that Hill's viewpoints could also be applied to individual studies and to portfolios of evidence. Thus these *criteria*, as they became known, provided an additional tool for assessing individual studies. They were developed before the use of RCTs became extensively used. Thus, Hill's criteria, which had been developed to assess epidemiologic or observational data, became used to inform RCT designs and to grade types of evidence.

The development of RCTs was given appreciable impetus by their potential ability to address most of Hill's criteria and simultaneously to compensate for confounders and uncertainties. However, the value of RCTs is often over-rated and they have limitations caused by their focused character, their dependence on the design study to deal with uncertainties and confounders, and their lack of generalizability to other sectors of the population that were not represented by the study group. Furthermore, good RCTs are expensive and difficult to apply to long-term outcomes such as those representing health and well-being and decreasing disease risk. They are also difficult to manage if they involve a sustained intervention and follow-up. In short, RCTs are a difficult tool for nutritional studies, unless, of course, the food component of interest has much in common with a pharmaceutical compound [5,6].

Irrespective of the types of evidence, if any portfolio has good quality evidence based on appropriate markers used in a mechanistic strategy derived from good quality studies (such as has been advocated in the FUFOSSE and in PASSCLAIM), Hill's criteria lend themselves to the creation and analysis of a well-structured argument for causality. Thus if the data presented in a portfolio essentially meet the elements of FUFOSSE and PASSCLAIM, then the completeness of the dataset can be gauged by the degree with which the information overall addresses the elements of Hill's criteria and maintains the integrity of the overall case in which the “totality of the evidence” is brought together to produce a coherent exposition of the case for establishing the extent of causality. The quality of the evidence is pivotal to this.

The flexibility of Hill's criteria to different types of evidence is becoming better appreciated. Importantly, proposed revisions of Hill's criteria are showing how the initial guidelines of experimental strength and temporality can be regarded as direct evidence, whereas those of biological gradient and biological plausibility are seen as mechanistic evidence; the criteria on coherence, consistency, and analogy can be seen to represent coherence, repeatability, and similarity as a group of parallel evidence [7]. Using this configuration, it can be shown how “evidence-based mechanistic reasoning” can be applied to demonstrating causality within the context of using Hill's criteria to assess the totality of the data [8]. These studies illustrate how a well-structured mechanistic pathway or process can be used to substantiate the extent to which causality can be demonstrated in the absence of RCTs or in the presence of weak RCTs. In the latter case, this adds weight to the criticisms of overdependence on RCTs in nutritional research and the need to use alternative forms of evidence of causality in nutrition [6]. The concept of evidence-based mechanistic reasoning is also consistent with the strategic use of markers for individual studies and the linking of independent studies proposed by FUFOSSE [1].

Another proposed benefit of using Hill's criteria was that these provided a framework within which allowance could be made for incomplete datasets and missing information. Statistical devices can be used to enable imputation approaches for missing data points, for the aggregation of incomplete datasets to enable “broken RCT analysis,” and for the statistical randomization of observational data [9]. Hill (1965) however warned against over precise use of analyses in the face of large systematic variation, and there may be inherent dangers in such manipulation of the data. Conversely, there is an increasing attention to the philosophy of using qualitative studies in exploring causality in the social sciences, education, and some aspects of health care [3,10,11].

Demonstrating causality: grading of the evidence

One of the problems associated with the assessment of scientific data underlying the relation between intake of a food or food component and an effect on health is how to deal with emerging science. By its very nature, emerging science is indicative of a health effect but needs further corroboration before it can be generally accepted as conclusive. Such corroboration can come from further controlled trials but also from other sources of evidence pointing toward the plausibility of the observed effect (e.g., experimental and animal studies supporting a plausible biological mechanism, observational evidence confirming the observed effect, etc.).

The World Health Organization (WHO) and the World Cancer Research Fund (WCRF) have applied a system of “grading” of the evidence as a scientific methodology to assess the amount

of evidence available and the strength and consistency of that evidence to support the plausibility of the observed effects [12, 13]. This system uses four grades of evidence: convincing, probable, possible, and insufficient. It has proved very helpful for public health institutions in the development of science-based public health recommendations.

Furthermore, if applied in the framework of claims assessment, it would enable regulators to take decisions on whether to accept an observed effect by balancing the strength of the claimed effect against the strength of the supportive evidence. Furthermore, it is crucial to support scientific research and product innovation to find an approach where the term *generally accepted scientific data* includes not only generic or well-established associations between a food or a food component and a health benefit but takes also into account the overall concept of the grades of evidence and the balance of probabilities that an association between a food or a food component and a health benefit will be refined (not reversed) by subsequent scientific research (Fig. 4).

Similar systems have been proposed (Appendix). In December 2004, the Netherlands Ministry of Health initiated discussions on the establishment of an inventory of substantiated health claims to fulfill the obligations of Article 13 of the European Union (EU) claims legislation then under development [14]. The basis of their proposed framework was a judgment and classification of the foods and food components and their health relations based on the strength and consistency of the scientific evidence in such a way as to underpin the definition of *generally accepted scientific data*. The approach developed the PASSCLAIM concept of a continuum of emerging and consensus science and used the

WHO/WCRF terminology to create five categories based on the grade of evidence:

Categories 1 and 2: insufficient substantiation; more data needed.

Category 3: probable/likely (possible); positive outweighs the negative evidence; balance of probabilities justifies the diet-health relation.

Categories 4 and 5: convincing; including evidence based on meta-analysis, peer-reviewed publications, text books, monographs, judgments by government-related organizations, scientific groups, or expert organizations (e.g., WHO, European Food Safety Authority, UK Scientific Advisory Committee on Nutrition).

This would enable assessors to identify generally accepted relations between foods and food components and health. It stresses the need to develop an acceptable procedure that takes into account the fact that scientific knowledge is constantly evolving and being refined.

Also, the European Medicines Agency proposed a system of grading for determining the requirements for scientific data underlying the various types of medicinal products [15].

Thus, the scientific grading of the evidence provides a way of expressing an assessment of the totality of the information available and enables a transparent scientific and qualified judgment relating to the strength, consistency, and biological plausibility of the relation between the intake of a food or food component and its role and effects for health. The best resource for assessing the totality of the evidence available in support of the plausibility of the beneficial effects of a specific food

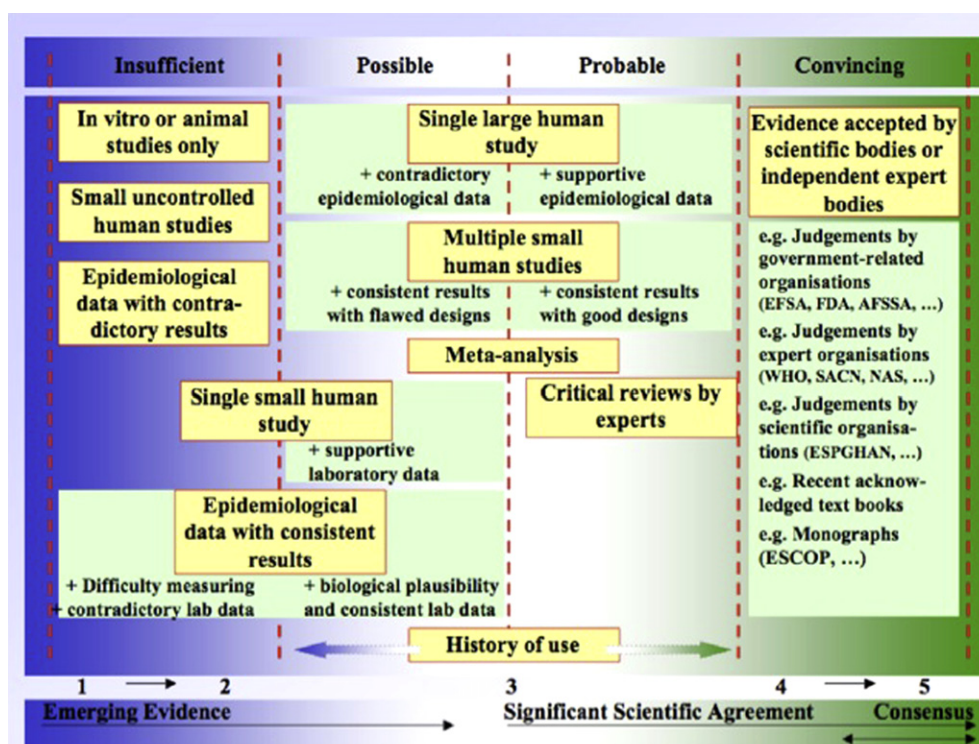


Fig. 4. Grading of the evidence. EFSA, European Food Safety Authority; FDA, Food and Drug Administration; AFSSA, Agence française de sécurité sanitaire des aliments; WHO, World Health Organization; SACN, Scientific Advisory Committee on Nutrition; NAS, National Academy of Sciences; ESPGHAN, European Society for Paediatric Gastroenterology; ESCOP, European Scientific Cooperative on Phytotherapy.

component is obviously an assessment by the leading experts in field. Such “consensus” meetings would be able to provide the best clarity on the current status of the scientific research, including its limitations and identify future research needs. It would also provide regulators with a sound foundation on which to base decisions for public health recommendations and the acceptance of health claims.

Assessing a claim

Assessment of the overall case for a claim is a risk-management process that should be taken by competent regulatory authorities. This agency would be provided with an assessment of the strength, consistency, and biological plausibility of the evidence for a health claim. Ideally this assessment would provide an explicit consideration of the weaknesses and strengths of the portfolio of evidence of causality and an account of the uncertainties and variabilities of the data as would be expected with a risk or scientific assessment.

Because the regulators will also need to consider the potential benefits and harms that might arise from their decision, the probability of these events would also need to be considered. It has been argued that assessments of degree of causality might be influenced by an awareness of the potential harm of not attributing or attributing causality. Thus, risk–benefit analysis might need to be a separate exercise from that of assessing the extent to which causality can be demonstrated [16].

The PASSCLAIM did not consider the qualification of claims should they be allowed. Various frameworks have been proposed, but the criteria for scientifically grading the strength of the evidence for causality have not been well defined. The grading of evidence in support of causality need not be the same as the grading of the contingent claim.

The competent authority would be expected to have other issues in mind when it decides whether to allow a claim. Such issues would include public health protection, public health improvement, enabling commercial freedom, and equitability in the market.

2. What is the process and what are the requirements to develop research methodologies that are able to demonstrate beneficial health effects on bodily functions in normal healthy populations?

Background

The primary goal of nutrition is to maintain, or if possible to improve, health. This is an essential difference from pharmaceuticals, which are generally developed to treat, cure, or prevent disease. The term *beneficial physiologic effect* refers to the demonstrable effect(s) of a food or food constituent for which a health claim is made. Beneficial physiologic effects go beyond traditionally accepted nutritional effects and their validity should be substantiated for the general population or a population subgroup. This requirement to demonstrate a beneficial physiologic effect(s) of a food or food constituent in the general healthy population presents an enormous scientific challenge to nutrition science. In fact, this task is more complicated than assessing the efficacy of most pharmaceuticals, for which effects on clinical or surrogate endpoints are often easier to establish.

The question on how to develop new research strategies will be split into the following subquestions: How to define and measure health? How healthy is the target population? How to

demonstrate beneficial physiologic effects in persons who are healthy?

What is health?

According to the WHO definition, health can be defined as “a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity” [17]. This definition takes into account that health may still be improved even if there are no obvious physical or mental conditions. However, it does not provide directions on how to quantify improvements in health, in particular not on a population scale. Therefore, it is proposed to use a definition that is based on the principles of homeostasis and at least one aspect of the dynamic of biological evolution: “health is the ability to adapt to internal and external stimuli” [18,19]. Homeostasis acts to maintain balance among the various biological processes that are interacting in an individual. The dynamics of these processes can be studied by looking at clusters of functional biomarkers, which in a healthy state are kept within a certain range. When there is long-lasting imbalance between the physical and mental robustness (adaptability) of an organism and external and internal stimuli, the risk for permanent harm (pathology) increases. The robustness of homeostasis at a particular moment can be determined using so-called challenge tests. It should be realized that health is time-dependent. Although a person can be healthy at a particular moment as shown by adaptability, there is a certain likelihood that genetic makeup and environmental factors will change this situation over time. This is called *risk for disease*. Currently, European legislation interprets beneficial physiologic effects in this case as to whether the food or constituent decreases the risk factor for the development of a human disease (not the risk of the disease itself). This interpretation can lead to complicated situations, e.g., the true outcome of a disease such as a cardiac event for coronary heart disease that is based on strong epidemiologic evidence and quantification of relative risk but no decreased risk factor from RCTs. These sorts of situations may occur, for example, with diseases for which no clear risk factor has been established or for which there is a lack of scientific consensus on the predictive value of such a risk factor. As a result, it might even happen that a food product that has shown to effectively decrease the risk for a particular disease can still be refused a health claim in the absence of an established and decreased risk factor. Clearly, there is a need for a review of the scientific implications of the absence of an established decreased risk factor in making claims for decreasing disease risk.

With regard to risk factors, there are several ways by which these may be modulated by diet or specific nutrients, e.g., at the molecular level. However, there is not much known yet about the interpretation of such interactions, in particular in the longer term. Therefore, it is recommended to adopt the establishment and validation of risk factors as a future research priority. Expert groups could be formed to reach greater consensus on the acceptability of potential risk factors. More fundamental research, e.g., in epigenetics, is needed on the interaction between genetic risk factors and nutrients. This discussion also identifies the option to abandon the “risk-factor” approach and refocus on biomarkers of optimal health based on the homeostasis concept. This is further discussed in section IV.

How healthy is the target population?

In many cases, the boundaries between health and disease are not well defined. Increasing knowledge on how diseases develop

and improved diagnostics have demonstrated that there is often a continuum between a healthy state and disease. At any given time, the general population consists of only a minor fraction that is “perfectly healthy,” whereas the majority is somewhere between an overt disease status and the “perfect-health” status. This holds particularly true for chronic diseases. For example, overweight (body mass index >25 kg/m²) and obesity (body mass index >30 kg/m²) currently affect more than 50 % of the adult population in the USA and Western Europe. Both conditions are important risk factors for diabetes type 2 and cardiovascular disease and contribute to the development of secondary complications. Persons who are overweight or obese often show increased markers for inflammation, deviant serum lipid profiles, increased liver enzyme values, and a slightly increased blood pressure without being diagnosed as patients. Therefore, to demonstrate beneficial physiologic effects, evidence in slightly impaired individuals may represent an alternative approach [18]. Clusters of biomarkers that reflect essential processes such as inflammation, oxidative stress, or metabolism can be used to construct a theoretical multidimensional “health space” [20]. This approach helps to reveal biological processes that may be functioning suboptimally and can be used to show more specific beneficial effects of intervention.

How to demonstrate beneficial physiologic effects in persons who are healthy?

Methods based on analysis of the robustness of physiologic homeostasis in individuals are currently regarded as the most promising approaches to measure health and beneficial physiologic effects of nutrition. Such approaches are generally following a challenge principle. Challenge tests should cover one or several relevant physiologic processes. Examples include variations of oral glucose and lipid tolerance tests, organ function tests, exercise, or even psychological stress challenges [21]. Although some of these tests, e.g., the glucose tolerance test, are not new, the combination with new bioanalytical technologies (microarray analysis and metabolomics) and calculation power makes them particularly useful to test health-improving effects of foods and food constituents. The designs of the tests should be optimized to include analyzing the time course of physiologic processes instead of single endpoints. In general, the intervention will take place after a crossover design, using a placebo if possible, with each participant being his/her own control. Although this will require intensive sampling and measurement schedules for each participant, this decreases the effect of interindividual variation and allows smaller numbers of participants per study. Challenge tests are promising for quantification of phenotypic changes and health effects of nutrition based on homeostatic adaptability, but there remain certainly issues to be resolved. These include the nature of the challenge, the force of the stimulus, accepted designs, statistics, and validation. Therefore, it is recommended to further develop and evaluate these approaches as concerted activities. As was also recently advised by the Proving the Efficacy of Foods and Food Constituents for Health Claims (PROCLAIM) taskforce, this should preferably be organized by international research consortia in which industry and academic groups work together [18].

II. Case studies

In this section several case studies were discussed to address under what conditions the totality of the current scientific data and

supportive evidence from generally accepted sources and recognized scientific institutions could find more suitable alternatives to RCTs.

These case studies address the following diet–health relations:

- Carotenoids and vitamin A in relation to AMD.
- The quality of carbohydrates (as expressed by the GI) in relation to health and well-being.
- Probiotics in relation to intestinal and immune functions.
- Micronutrient intake and maintenance of normal body functions.
- Food components with antioxidative properties and health benefits.
- The nature of evidence supporting the impact of deficient, adequate, and optimal intakes of micronutrients on physiologic function.

The case studies were chosen as examples to address several questions, including:

- How far is an evidence-based benefit a reliable endpoint?
- To what extent are the data and practical experience of health care professionals capable of providing additional scientific support?
- To what extent can data from clinical studies of disease states be used as supportive evidence for health effects?
- How to assess the different effects of the various nutritional components from the diet on systemic parameters, e.g., the individual effects of the various types of fatty acids in the diet on blood lipids, the individual effects of the various types of carbohydrates and carbohydrate-containing foods on metabolic and glycemic responses, and in particular the effects of their mutual replacement in the diet.
- How to identify the value of changes in biological parameters of complex physiologic systems (e.g., immune function, metabolic responses) for the identification of beneficial effects.

Consensus statements

Practice and experience by health care professionals can be well established but mostly are not documented and may therefore not be part of the evidence currently required for a claims assessment. However, these form part of the evidence acknowledged by the experts in the field. If because of this situation claims are rejected, consumers are denied the use of information on products, despite their physicians or other health care workers advising them.

Although the review process of health claims has identified the weaknesses of current nutritional research (and the peer-review process), the acquired scientific knowledge available to date cannot be ignored and should be a starting point for the assessment of the totality of the available data and the strength, consistency, and biological plausibility of the evidence and for the identification of stronger research concepts.

There is a need to point out that the focus on the presence of a decreased risk factor is a limiting factor and that foods and food constituents can have an effect on the decrease of a disease risk, although a risk factor cannot be identified. Nutritional risk factors should also be acceptable.

There are specific areas that appear not to fit the legal framework for the assessment of the claims, such as GI claims (as defined by the claimed effect) and botanicals (based on traditional use). Instead of letting these established health effects fall out of the framework, the regulatory framework should recognize the nature and source of this evidence base and be modified to fit them in.

There is a need to translate the impact of nutrition on health and wellness (e.g., cognitive performance, inflammatory processes, cardiovascular disease risk, etc.) into a framework of biomarkers and physiologic responses sufficient to quantify and substantiate these relations for the maintenance of a beneficial bodily function.

1. Carotenoids and vitamin A in relation to AMD

As far back as the ancient Egyptians, it was appreciated that night blindness could be treated with dietary interventions such as the consumption of liver. Millennia later, it was demonstrated through biochemical and animal studies that the liver is rich in vitamin A whose metabolite, 11-*cis*-retinal, is critical for the normal function of the rhodopsin photopigment in the human retina. It also became clear that a non-essential nutrient precursor of vitamin A, β -carotene, could substitute for vitamin A in a manner that posed much less risk of toxic side effects. The effects of β -carotene are so dramatic and clear-cut that it has become widely accepted as a valid public health intervention against eye diseases related to vitamin A deficiency in the developing world without a large-scale RCT to support such a claim.

Age-related macular degeneration is the leading cause of blindness in people older than 50 y in the developed world [22]. The number of adults registered blind as a result of AMD in industrialized countries continues to increase, primarily due to increasing longevity [23]. Beyond its inevitable impact on the individual, AMD poses a growing socioeconomic challenge to modern society [24]. To determine the role of nutrients in the maintenance of eye health is a challenging problem. AMD is a slowly progressive complex disease influenced by genetics, aging, and multiple environmental factors including smoking and light exposure [25]. Epidemiologic studies of the 1980s indicated that diets rich in carotenoids could be a mitigating factor for the risk of AMD. So, when the Age-Related Eye Disease Study (AREDS) RCT was conceived, the best known and most widely available carotenoid, β -carotene, was included with zinc, vitamin C, and vitamin E. The positive AREDS results published in 2001 have had a profound impact on clinical retina practice in the USA because AREDS-type supplements are routinely prescribed to patients with the intermediate to advanced stages of AMD, but the recommendations were not without controversy [26].

There was no guidance as to whether AREDS supplements would be helpful in patients with earlier stages of AMD or with a strong family history. Because β -carotene, vitamin C, and vitamin E were given together, it was not clear which nutrients were actually driving the positive effects. Two other studies published while AREDS was in progress found that high-dose β -carotene supplementation significantly increases risk of lung cancer in smokers [27,28]. Also, β -carotene itself is not found in the retina, and its metabolism to vitamin A is tightly regulated. As such, it became increasingly recognized that other carotenoids such as lutein and zeaxanthin, which are found in the macula at high concentrations, might have been better choices than β -carotene for intervention against AMD.

Indeed, after the inception of AREDS, it became clear from the many epidemiologic and biochemical studies that lutein and zeaxanthin, rather than β -carotene, are likely to be the protective factors in carotenoid-rich fruits and vegetables [29]. Accordingly, the AREDS2 study is testing lutein 10 mg and zeaxanthin 2 mg in a randomized, placebo-controlled, double-blind clinical trial in addition to ω -3 polyunsaturated fatty acids 1000 mg derived from fish oil.

The retinal carotenoids lutein, zeaxanthin, and *meso*-zeaxanthin accumulate in the macula, where they are collectively referred to as

macular pigment [30]. Lutein and zeaxanthin are of dietary origin, whereas *meso*-zeaxanthin is not normally found in a conventional diet and is thought to be generated at the retina after lutein isomerization [31]. In brief, macular pigment is a short-wavelength (blue) light filter and a powerful antioxidant, and it is therefore believed to protect against AMD. Although the concept that macular pigment protects against AMD remains a hypothesis, the rationale in support of this view is biologically plausible and is supported by the findings of in vitro and animal studies in which lutein, zeaxanthin, and *meso*-zeaxanthin have been shown to protect photoreceptors against oxidative injury [32].

Currently, eye care specialists (e.g., ophthalmologists and optometrists) find themselves in a difficult position when attempting to make sound and evidence-based recommendations to patients at risk for visual loss from AMD. The AREDS mixture remains the only formulation that has been shown to be of benefit in AMD in the context of a well-designed RCT involving a large number of patients ($n = 4531$). However, the AREDS formulation contained β -carotene (discussed earlier), and the amount of zinc in the AREDS formulation exceeded the EU upper safety limits [33]. However, it is difficult to ignore the basic implication of AREDS, namely that antioxidants are beneficial for patients with AMD. Such an interpretation has encouraged the nutraceutical industry to promote the use of antioxidant supplements that do not include β -carotene, are EU compliant, and contain the macular carotenoids. In 2004, the Lutein Antioxidant Supplementation Trial (LAST) was undertaken to investigate whether nutritional supplementation with lutein alone or lutein plus antioxidants, vitamins, and minerals improved or stabilized visual function in 90 patients with advanced atrophic AMD [34]. This study reported that visual function improved with lutein supplementation alone or with lutein plus a combination of antioxidants, vitamins, and minerals compared with a control.

One may understand why an eye care specialist, in the absence of convincing evidence from RCTs but in the presence of a biologically plausible rationale and significant body of supporting evidence (albeit not from RCTs), might recommend antioxidant supplements that contain the macular carotenoids in view of the lack of other available putative or proved measures against AMD. The patient with intermediate or advanced AMD, who may have already lost vision in one eye, often wishes to actively participate in decreasing the risk against further visual loss and is unwilling to wait for a “conclusive” evidence base. Similarly, the “worried well” who have early signs of AMD or who have a strong family history of AMD are unlikely to be willing to wait for definitive RCTs applicable to them, especially in light of the excellent safety record of carotenoids when administered at the AREDS2 dose. Nevertheless, under these circumstances, it is incumbent on the eye care professional to inform patients with AMD that such supplements have not been proved to protect against the development or progression of AMD but have strong indications of benefit.

In conclusion, while we await the outcomes of AREDS2 and other ongoing RCTs, the potential public health benefits, scientific plausibility, and excellent safety records of lutein, zeaxanthin, and *meso*-zeaxanthin are supporting the recommendation that dietary consumption and supplementation with these compounds should be encouraged. In the absence of definitive RCTs, we must view and assess the totality of the scientific data and weight of evidence as currently available, identifying and acknowledging such evidence from respected sources and scientific institutions, to find more suitable alternatives to RCTs to make important public health recommendations in a timely manner.

2. Quality of carbohydrates (as expressed by the GI) in relation to health and well-being

Food chemists have typically categorized dietary carbohydrates into simple sugars and complex carbohydrates on the basis of their degree of polymerization. This form of classification of carbohydrates is a well-established concept in food science. However, the effect of carbohydrate on health may be better categorized according to their physiologic effects, notably their ability to increase blood glucose. The blood glucose response varies substantially among different carbohydrate-containing foods and cannot be predicted by their gross chemical composition alone. This gave rise to the concept of the GI.

The GI was developed by Jenkins et al. [35] in 1981 as an improvement of the carbohydrate exchange list for diabetics. It is a numerical ranking of carbohydrate foods defined as the incremental area under the blood glucose response curve of a 50-g carbohydrate portion of a test food expressed as a percentage of the response to the same amount of carbohydrate from a standard food (usually white bread or glucose) taken by the same subject. By this definition, carbohydrate foods are rated as low (<55), medium (55–70), and high (>70) GI foods.

A series of studies have confirmed that consuming a low GI diet (<55) has a significant effect on decreasing hemoglobin A1c and fructosamine, confirming that the widely used cutoff points for GI classification (<55, 55–70, and >70) have an identifiable clinical outcome [36]. Emerging data also suggest that the postprandial glycemic spike and the degree of blood glucose excursion may also be of clinical importance. Foods that show smaller peak changes in glycemia and maximum amplitude of the glucose excursion provide greater metabolic benefits to human subjects (Table 1) [37].

The role of fasting blood glucose in the development of coronary heart disease is less well recognized, mainly because of the paucity of studies correlating fasting blood glucose to coronary heart disease. Bjørnholt et al. [38] reported the results from a 22-y prospective study on fasting blood glucose as a predictor of cardiovascular death. Of 1990 apparently healthy non-diabetic men (40–59 y old), 1973 with fasting blood glucose levels lower than 6.1 mmol/L (110 mg/dL) were included in the study. After 22 y of follow-up, 483 men had died, 53% from cardiovascular diseases. It was also found that men in the highest glucose

quartile (fasting blood glucose >4.7 mmol/L [85 mg/dL]) had a significantly higher mortality rate from cardiovascular disease compared with those in the lowest three quartiles (Fig. 5A–C).

Another study by the Asia Pacific Cohort Studies Collaboration [39] collated fasting blood glucose data from 237468 participants; during approximately 1.2 million person-years of follow-up, there were 1661 stroke and 816 ischemic heart disease events [39]. Continuous positive associations were demonstrated between usual fasting glucose (≥ 4.9 mmol/L) and the risks of

Table 1
Association among GI, peak blood glucose changes, and MAGE [37]

Food category and GI category	GI	Peak change (mmol/L)	MAGE (mmol/L)
White breads			
Low GI	49 ± 1	1.81 ± 0.06	1.82 ± 0.06
Medium GI	63 ± 1	2.20 ± 0.06	2.21 ± 0.06
High GI	76 ± 1	2.45 ± 0.04	2.49 ± 0.05
P	<0.001	<0.001	<0.001
Whole-grain breads			
Low GI	49 ± 1	1.95 ± 0.04	2.05 ± 0.04
Medium GI	62 ± 1	2.21 ± 0.03	2.24 ± 0.04
High GI	74 ± 1	2.29 ± 0.05	2.29 ± 0.05
P	<0.001	<0.001	<0.005
Breakfast cereals			
Low GI	46 ± 1	2.10 ± 0.08	2.18 ± 0.09
Medium GI	63 ± 1	2.60 ± 0.08	2.78 ± 0.08
High GI	81 ± 2	2.89 ± 0.08	3.02 ± 0.08
P	<0.001	<0.001	<0.001
Cereal snack bars			
Low GI	43 ± 1	2.00 ± 0.008	2.14 ± 0.08
Medium GI	62 ± 1	2.66 ± 0.09	2.82 ± 0.09
High GI	77 ± 2	2.93 ± 0.11	3.07 ± 0.10
P	<0.001	<0.001	<0.001

GI, glycemic index; MAGE, maximum amplitude of glucose excursion

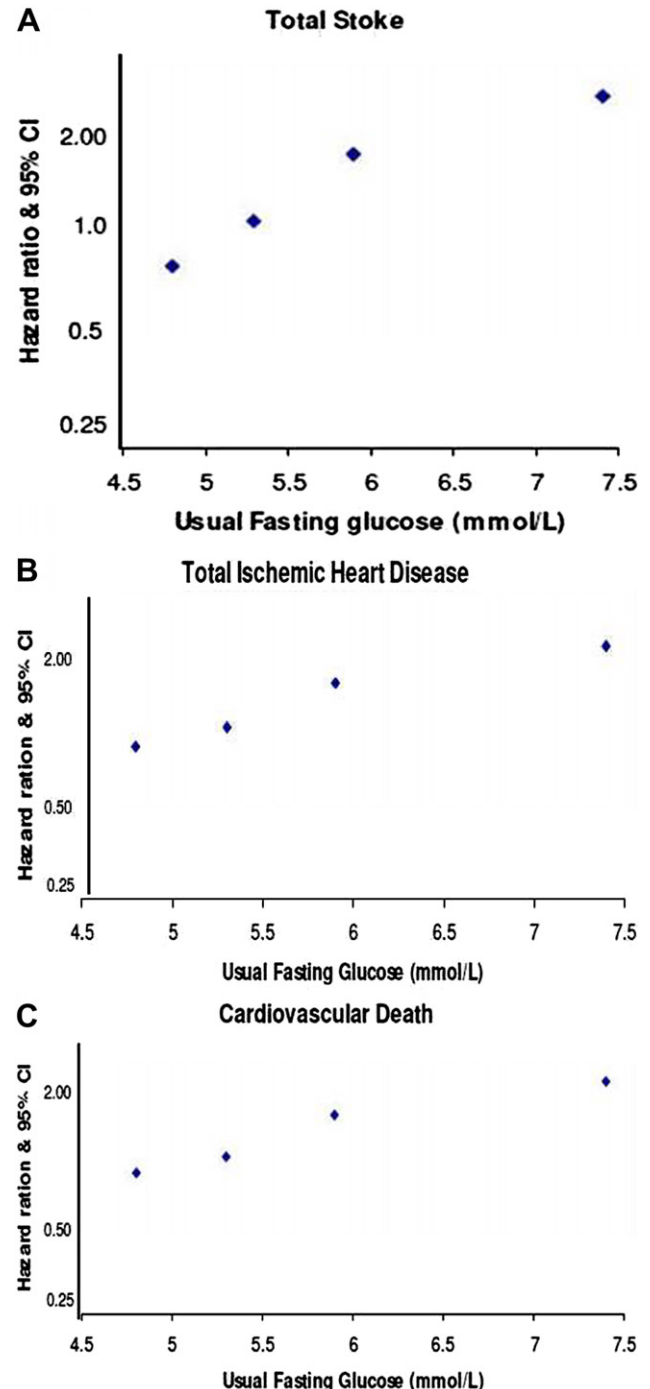


Fig. 5. The relation between fasting blood glucose and (A) stroke, (B) total ischemic heart disease, and (C) cardiovascular death. Adapted from the Asia Pacific Cohort Studies Collaboration [39]. CI, confidence interval.

cardiovascular disease. Each 1-mmol/L decrease in usual fasting glucose was associated with a 21% (95% confidence interval 18–24) lower risk of total stroke and a 23% (19–27) lower risk of total ischemic heart disease. The associations were similar in men and women, across age groups, and in Asian compared with Australasian (Australia and New Zealand) populations. Fasting blood glucose is an important determinant of cardiovascular disease burden, with considerable potential benefit of usual blood glucose lowering down to levels of at least 4.9 mmol/L.

For several decades the diet–chronic disease paradigm has centered on decreasing the intake of saturated fat and cholesterol and increasing the consumption of complex carbohydrates. The type of carbohydrate consumed may be as or more important than the type of fat consumed: It is here that the concept of the GI may have a significant role to play [36]. The increased intake of refined carbohydrates (i.e., high GI carbohydrates) in most Western diets may have aggravated the incidence of cardiovascular disease and diabetes.

3. Probiotics in relation to intestinal and immune functions

Humans are surrounded by a world of micro-organisms. The immune system is a complex physiologic system, the function of which is to protect the individual against infectious diseases that otherwise could be caused by those micro-organisms. Cells of the immune system, by virtue of antigen-specific receptors, can recognize non-self from self and thus have the ability to react to invading micro-organisms. Along the same self/non-self principle, the immune system also recognizes and destroys transformed cells, thus protecting against malignancies. Close regulation of the immune system is necessary because an overactive immune system may lead to chronic inflammatory diseases, allergic diseases, or autoimmune diseases.

The human immune system operates at two functional levels: innate and acquired immunity. The innate responses are activated within hours of encountering an antigen, whereas the specific adaptive immunity takes days or weeks to develop and act. Phagocytes and natural killer cells produce cytokines and chemokines that together with complements form the innate responses. Innate immunity, in particular the dendritic cells, initiate and orchestrate the adaptive immune system, which consists of B and T lymphocytes. The hallmark of acquired immunity is antigen specificity and immunologic memory. Under normal physiologic conditions, the immune system protects the individual against most pathogens. The function of the immune system can be temporarily impaired because of intrinsic or extrinsic factors, permanently weakened in the elderly, and is not fully mature during infancy. Even short periods of impaired immunity can compromise the ability to fight infections.

Albers et al. [40] identified 28 different groups of markers to measure immunomodulation in human nutrition intervention studies. Immunoglobulin (Ig) levels in serum and at mucosal surfaces, systemic cytokine concentrations, number and activity of phagocytic cells, and natural killer cells are among the biomarkers most often studied. None of these biomarkers by themselves are indicative for the overall immune status of an individual. The number of CD4⁺ T lymphocytes in peripheral blood is an excellent marker for the immune status of patients infected with the human immunodeficiency virus and to monitor immune reconstitution after therapy. For healthy individuals with CD4 counts within the normal range, this parameter has limited value as a biomarker for the status of the cellular immune system. The same holds true for IgA. This class of antibodies is especially designed to function at mucosal surfaces of the respiratory and digestive tracts. The mean

serum IgA concentration in health adults is 2.08 g/L (interquartile range 1.49–2.68) [41]. Selective IgA deficiency, defined as a serum IgA concentration below 0.05 g/L with normal levels of IgG and IgM, is the most common form of humoral immunodeficiency, occurring in about 1:700 individuals. Patients may have severe and prolonged gastrointestinal and respiratory infections, but about half of individuals with selective IgA deficiency are healthy [42]. Against this background, it is difficult to predict, e.g., whether a 25% increase in the concentration of secretory IgA in the mucosal surfaces of the upper respiratory tract really would confer better protection against influenza virus infection. This is the reason claims such as “strengthens the immune system” are considered vague or even meaningless, unless they relate to a more specific claimed health effect. There is a gap between the scientific understanding of the immune markers and immune system and the regulatory requirement for measuring the specific effects on immune health [43].

The best biomarker for the status of the immune system would be the clinical response to infection and/or the serologic response to vaccination.

To demonstrate that probiotics can actually decrease the risk for gastrointestinal and respiratory infections, many clinical trials have been conducted [44]. In most cases, the study population consisted of participants with an increased risk of infection, namely the (very) young and the (very) old. The frequency of respiratory tract infections in otherwise healthy toddlers in day-care centers and the frequency of flulike infections during the winter season in otherwise healthy elderly have been shown to be decreased by probiotics [45]. To demonstrate the causality of the probiotic effect on decreasing disease risk, changes in the components of the immune system that contribute to protection against a rhinovirus infection have been determined [45]. For future clinical studies, therefore, biomarkers of the immune system should be included to demonstrate the causality of the beneficial effect [43].

The downside of clinical trials in which incidence, duration, and severity of infectious episodes are recorded is that in most cases the causative microbiological agent is not identified. This hampers the selection of the most appropriate biomarker of the immune system. An alternative approach to establish a cause-and-effect relation is to deliberately infect humans with a (mild) pathogen. This has been done for enterotoxigenic *E. coli*, *Staphylococcus aureus*, and rhinovirus [46]. Such trials are costly and difficult to set up. A more feasible approach to measure the overall effect of a nutritional intervention on the in vivo function of the immune system is to measure the response to vaccination [47]. The two vaccines that would be most appropriate for this purpose are influenza and pneumococcal vaccines. In most European countries, seasonal influenza vaccine is administered yearly to elderly and other risk groups. It is known that in the elderly the response to vaccination is suboptimal (and therefore provides incomplete protection). Pneumococcal polysaccharide conjugate vaccines are incorporated in most childhood vaccination schedules. Pneumococcal polysaccharide vaccines are recommended for persons older than 65 y. For influenza and pneumococcal antibodies, assays for quantification have been standardized and the antibody concentrations that serve as correlates for protection have been established [48,49].

The serologic response to vaccination measures the response of the humoral immune system. An in vivo effector function of the cellular immune system is the so-called delayed type hypersensitivity reaction. As yet there is limited experience with delayed type hypersensitivity as a biomarker for cellular immune status in nutritional intervention studies.

The dynamic changes in systemic antibody concentrations after vaccination can be used for the identification of beneficial effects (of probiotics or other food components/additives) on the immune system. Validated antibody concentrations that serve as correlates of protection exist for the influenza and pneumococcal vaccinations and should be used for that purpose.

4. What is the nature of evidence supporting the impact of deficient, adequate, and optimal intakes of micronutrients on physiologic function?

Nutritional status and physiologic responses are intimately linked but most often lack clearly defined, specific, and validated functional tests necessary to demonstrate this relation in a population of generally healthy people. The absence of such assessments hampers our ability to develop new nutrient requirements and promulgate better dietary guidelines.

Employing a classic nutritional paradigm, inadequate intakes of essential nutrients are defined by the development of clinical signs of a well-characterized deficiency syndrome. Depending on the biological half-life of the specific nutrient, the symptoms of the deficiency will be manifest as short latency diseases within months, e.g., rickets from vitamin D deficiency or night blindness from vitamin A deficiency. The acute solution to this problem is the simple intake of the nutrient and resolution of the deficiency. In contrast, a low intake or higher biological demands of a nutrient over a period of years may result in declining physiological function and the development of a long-latency chronic disease. The solution to this problem is more complex and requires the use of biomarkers and/or other measurements of assessing the slowing or reversal of the rate of the specific pathophysiology. How then can we define adequate or optimal nutrient intakes to promote health or decrease the risk of a chronic disease?

To calculate the adequate intake of a nutrient sufficient to prevent a deficiency syndrome, different reference values have been employed. For example, the *estimated average requirement* is defined as an average daily intake for healthy individuals that achieve an established functional or clinical endpoint of adequacy. The estimated average requirement is intended to reflect the median for a population, i.e., half the population should fall below and half above this value.

Importantly, the estimated average requirements listed in Table 2 can be markedly lower than the recommended dietary allowance and the daily value used in nutritional labeling [50]. To specify the adequacy of intake, biochemical parameters such as plasma or tissue status have often been employed. However, this approach is often difficult to accomplish in large populations and often does not accurately reflect clinical or functional outcomes.

Thus, an estimated average requirement might better be defined using a static biomarker, such as plasma concentration, and a functional measurement. For several micronutrients, including vitamins A and E and the minerals iron and zinc, static indicators

are not available to define an adequate (or optimal) intake (Table 3) [51]. Furthermore, this approach is currently limited to the assessment of the adequacy of essential nutrients and is not readily applicable to other dietary constituents such as phytochemicals or fatty acids other than α -linolenic and linoleic acid.

There are few good examples of establishing optimal requirements of vitamins or other nutrients for long-latency chronic diseases. Efforts have been undertaken to evaluate optimal intakes of vitamin D by evaluating the results of observational studies and randomized clinical trials. Very recently, the Institute of Medicine (IOM) presented new dietary reference data for calcium and vitamin D [52]. The data for vitamin D were based primarily on the intake and 25(OH)-D serum concentration needed to maintain healthy bones. The IOM panel did not find scientific evidence for further beneficial effect of an adequate vitamin D intake on non-skeletal benefits (for review, see Whiting and Calvo [51]).

5. Food components with antioxidative properties and health benefits

Many antioxidants exist in food and might exert their antioxidant effects after ingestion with a meal within a complex pattern of further known and unknown antioxidants or as an isolated and specified antioxidant component in a food or food supplement. Related to their nature as antioxidants, nearly all are claimed to decrease reactive oxygen species (ROS) and subsequently decrease oxidative damage and compensate oxidative stress. A major paradigm that justifies the use of antioxidants is that ROS exposure of cells and tissues results in oxidative damage, impairment of function, and at least a disease state. It has been shown in some studies that ROS are responsible for the development of different diseases such as cancer, neurologic diseases, coronary heart disease, and other chronic degenerative diseases [53]. The major question with respect to antioxidants and their “evidence of benefit” is whether the detection of an impact of antioxidants on oxidative stress might be sufficient to claim a beneficial effect or whether a combination of a disease-related marker and the antioxidative effect is needed. Different analytical methods exist to measure the antioxidative activity in humans. Because of the distribution of different ROS in different cellular compartments, it is necessary to combine these methods to detect intra- and extracellular ROS. In cases of degenerative disease, the time from oxidative damage to occurrence of

Table 2
Examples of EAR [50]

Nutrient	Daily value	Highest RDA	Population-weighted EAR
Vitamin A	5000 IU	900 RAE	529 RAE
Folate (μ g)	400	400	314
Calcium (mg)	1000	1300	1091*
Iron (mg)	18	18	6.1
Zinc (mg)	15	11	7.5

EAR, estimated average requirement; RAE, retinol activity equivalents (μ g); RDA, recommended dietary allowance

* Based on adequate intake.

Table 3
Examples of indicators used to set estimated average requirements for vitamins (adult values) [51]

Nutrient	Static indicator of transport or stores	Indicator of function
Vitamin C	near maximal leucocyte concentration	
Vitamin E		prevention of H ₂ O ₂ -induced hemolysis
Folate	erythrocyte folate, serum folate	plasma homocysteine
Vitamin B12	serum B12 status	maintenance of hematologic status
Vitamin B6	plasma pyridoxal-phosphate status (women)	urinary xanthurenic acid after a tryptophan load (men); plasma homocysteine concentration
Thiamin	urinary thiamine	erythrocyte transketolase
Riboflavin	urinary riboflavin	erythrocyte glutathione reductase
Niacin	urinary methyl-nicotinamide	
Vitamin A	adequate hepatic vitamin A stores	

a disease-related biomarker is too long (years) to design a realistic human intervention study. If there is a relation between oxidative damage and disease development, a documented decrease of oxidative damage would be evidence of benefit similar to the accepted evidence for the benefit of decrease of low-density lipoprotein cholesterol. The development of a degenerative disease is a long process and not triggered by one effect alone. The risk for coronary heart disease increases with high cholesterol or triacylglycerols, obesity, increased ROS formation, or a diet with a low concentration of antioxidants.

Dietary intakes, in particular of antioxidants, are a key determinant of gene expression, in part by their involvement in genomic stability. Long-lasting degenerative diseases, such as cancer, coronary heart disease, chronic inflammation, and accelerated aging, appear to be caused in part by damaged DNA. DNA damage can be determined by measurement of micronuclei (MN) base lesions, strand breaks, and hypo- or hypermethylation. There are data from basic research and intriguing evidence from human studies to suggest that insufficient intake of vitamins A, B6, B12, E, folate, and zinc and phytochemicals, such as some carotenoids and flavonoids, can affect the rate of spontaneous chromosome damage. For example, to determine the effect of nutrient intake on DNA stability, Fenech et al. [54] measured the percentage change of MN in lymphocytes of 190 human volunteers and found several vitamins and minerals that were able to significantly protect DNA from damage and/or promoted DNA repair (Fig. 6).

Thus, changes in MN frequency might be used in intervention studies to evaluate the impact of different levels of micronutrient intake. Clinical trials have already shown that plasma statuses of folate and vitamin B12 are inversely associated with the degree of DNA hypomethylation, chromosome breaks, and MN formation [55,56]. The formation of MN in dividing (binuclear) cells is a result of chromosome breakage or malsegregation, excessive uracil incorporation into DNA, or hypomethylation of centromeric DNA [57]. Increased frequency of MN has also been demonstrated in smokers as a result of strand breaks or base lesions caused by oxidative stress associated with an inadequate intake of antioxidants [58]. Further, a lack of sufficient magnesium or calcium, minerals required for kinetochore or spindle assembly and DNA repair, can also increase MN formation [59]. Because MN formation changes in

relation to micronutrient intake, this method is as a sensitive functional indicator for the optimal intake of micronutrients necessary for genome stability.

Thus, functional indices of immune responsiveness and DNA damage may be useful biomarkers for defining adequate and optimal intakes of some micronutrients. These methods are particularly suitable for evaluating in relatively small-scale studies and, when associated with static marker or other surrogate endpoints, can provide valuable information on the physiologic and metabolic mechanisms related to adequate and optimal intakes and the risk of long-latency chronic disease.

6. How to assess the effects and health benefit of botanicals as food components?

Much of what we know about the health effects of botanicals and botanical preparations stems from knowledge and experience that humans have accumulated over time [60]. The properties of botanicals that were part of their natural environment were carefully observed when applied and led to the optimization of their use within a particular society, and this process of observation occurred long before the arrival of clinical trials and sophisticated analytical methods.

In Europe, medicinal law has accepted this observational approach and provides for a simplified registration based on traditional use for medicinal products [61]. Food law, however, expects RCTs to demonstrate health effects [62]. Such data are not available in most cases for botanicals and the regulatory process risks eradicating much of the heritage of acquired knowledge.

Furthermore, addressing the beneficial effects of botanicals and botanical components on health using RCTs presents challenges that cannot in the current state of science be solved.

- An approach to assess the efficacy of individual molecular substances is not adapted to be applied to complex mixtures of naturally occurring substances as present in botanicals. For single molecules, the pharmacokinetics and pharmacodynamics can easily be studied and clinical trials easily set up; for the complex mixtures that botanicals are, it is nearly impossible to use the same methodology [63,64].
- The demonstration of a physiologic effect in healthy people by intervention trials is almost impossible because it concerns an effect beneficial to the maintenance of a healthy

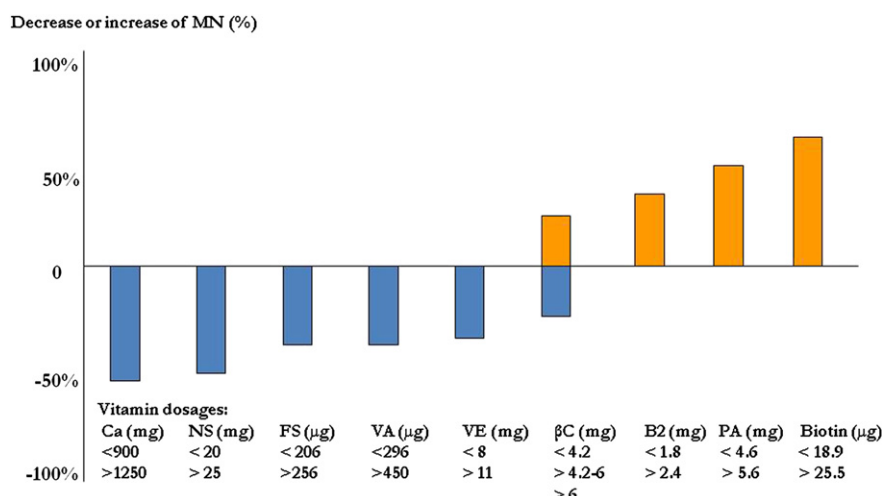


Fig. 6. Decrease or increase (percentage) of MN due to different dosages of vitamins [54]. βC, β-carotene; Ca, calcium; FS, folic acid; MN, micronuclei; NS, nicotinic acid; PA, pantothenic acid; VA, vitamin A; VE, vitamin E.

normal state (homeostasis) and to people not having symptoms of pathologies [65]. It is therefore nearly impossible to determine and assess adequate markers and to define these with precision in healthy people.

- Effects relating to the maintenance of a health status often become obvious over a long period (sometimes years or even decades). It is not feasible to assess such effects through intervention trials. Observational evidence and knowledge acquired over time are valid elements and therefore part of the evidence to help judge the strength, plausibility, and consistency of a certain effect on health [66,67].
- New techniques such as metabolomics are emerging and could perhaps solve in the distant future some problems linked to the use of complex systems such as botanicals [68]. However, this technique is extremely expensive, needs chemically very well-defined extracts, and the outcomes will probably not be valid in the foreseeable future until sufficient reference data have been built up [69].

The use of the vast body of empirical and traditional knowledge available for many botanicals and their preparations is therefore essential for the appreciation of their beneficial effects. This volume of traditional and empirical knowledge is gradually being confirmed by fundamental and applied research and reported in the scientific literature because of the current possibilities of phytochemical research and pharmaco-toxicological study of the components of these plants.

The European nutrition and health claims legislation has formalized the approval of health claims on foods, and the traditional use of many botanicals now falls within this framework. They all risk being rejected because of the reliance on RCTs, which are not available for this group of products. Although for decades botanicals have been accepted on the basis of experience, a new approach needs to be developed to enable a judgment on the validity of traditional use for botanicals. It is therefore proposed that, for the purpose of claims assessment, documentation needs to be provided to demonstrate traditional use of a botanical or botanical preparation for a certain health effect and all possible sources of information should be considered. The more evidence there is from these various sources, the stronger is the support for an acceptable tradition of use.

- Health effects of botanicals and botanical preparations have been recorded over time in historic and ethnobotanical records, botanical reference textbooks, monographs, etc. The more extensive these records describe the observed effect, the stronger the support for a traditional effect.
- It must be clear that the botanical preparation for which traditional information is used must be in line with the traditional form (e.g., infusions, aqueous or hydroalcoholic extracts, or their dried form). The more a product deviates from its traditionally used form and conditions of use, the less certainty there is about the validity of the traditionally observed physiologic effect.
- There are multiple recent reviews or compilations of traditional health effects in various monographs (WHO, European Scientific Cooperation on Phytotherapy, The Committee on Herbal Medicinal Products, Pharmacopoeia, Kommission E, leading experts in this field, etc.) that are generally recognized as authoritative and can be used to help substantiate the effect [70].
- Further evidence can be given by more recent experimental and applied scientific data that confirm the effect or are indicative of the underlying biological mechanism.

Pharmacologic data from chemical classes to which the known components of the botanical or the preparation belong may also support the plausibility of the traditional effect [71,72].

Once this information is collected, a good overview on the traditional and scientific knowledge of a botanical or botanical preparation will be available and allow an assessment of the validity and strength of the traditional claim.

III. Role of various types of scientific data: clinical trials, meta-analysis, and observational and traditional evidence

One particular field of discussion is the value of the various sources of scientific data in support of a diet–health relation. This includes the advantages and limitations of sources including RCTs, meta-analysis, observational evidence, and traditional use in the process of assessing the evidence for diet–health relations. This section addresses these aspects of the claims assessment process.

Consensus statement

There is a need to propose criteria for the identification of pertinent data and identify methodologies available to address the totality of this evidence according to the strength of the various study designs and results (as is applied to the assessment for diet–health relations as the basis for public health recommendations).

There is a need to consider the “inverse precautionary principle”: if the evidence is not absolutely conclusive but substantially indicative of the effect, then take a management decision to allow the claimed effect for the benefit of the consumer. An assessment of the evidence should contain the necessary information to make such management decisions possible.

The evidence needed should be proportionate to the strength of the claim.

1. Are RCTs capable of addressing the complex interactions of dietary interventions and of isolating the effects of individual food components within the food matrix?

In section I, it was concluded that experience suggests that current assessments depend too much on evidence from RCTs, which was not PASSCLAIM's intention. Furthermore, RCTs, whether “explanatory” or “pragmatic,” are not always possible or appropriate.

The presence or absence of positive RCT data should not therefore be a gate-keeper criterion in the evaluation of nutrient or food claims for several reasons.

- For applications involving new indications for a recognized essential nutrient, an adequate RCT would not be ethically permissible because the control group would have to be placed on an intake low enough to result in the untoward outcomes that an adequate intake would have prevented. Prospective cohort studies, case–control studies, observational studies of high quality, crossover studies, and experimental data can contribute to the totality of evidence in cases when a RCT is impossible or unethical.
- Even when a RCT would be ethically permissible, the capability of RCT designs for foods or nutrients is markedly limited for the control of critical effect modifiers such as synergies with a multitude of other nutrients and the influence of nutrigenomic relations.
- Because of the ubiquity of nutrients, for most nutritional interventions tested in a RCT, a true placebo or control group is often

not possible because no dietary pattern can feasibly exclude exposure to the test nutrient before and/or during the trial.

- Where validated intermediary biomarkers are not available, the duration of a RCT to test the ability of a nutritional intervention for the primary prevention of a chronic disease may require a decade or more and thousands of subjects. Thus, if required as a necessary criterion for evidence of a health benefit, RCTs often do not present a feasible research approach. Their being required impedes innovations and applications by the food and nutritional industries.
- Society does not need the same level of certainty concerning the effect of a low-risk nutrient or food as it does for a potentially higher-risk (and much more expensive) drug. For example, if the input/exposure concerned is lower than the amount listed as safe by responsible authorities (e.g., lutein) and if observational studies indicate the likelihood of a favorable effect, such a claim could be accepted because the harm of not accepting (i.e., possible benefits forgone) would outweigh the harm of approval (which would appear to be negligible even if the effect is actually null). Each case needs to be evaluated individually and on its own merits, but the absence of RCT evidence in such cases should not preclude evaluation of a submission. In addition, to require for a low-risk nutrient the assurance appropriately necessary for a new drug is likely to stifle product innovation. What has been termed the *inverse precautionary principle* may be helpful in this context. If the evidence is not absolutely conclusive but substantially indicative of the effect, then a management decision to allow the claimed effect may be taken for the benefit of the consumer.

Although it is certainly true that the RCT is the only trial design that supports strong causal inference, this statement is correct only if the RCT meets several criteria.

- Generalizability to the target population.
- Minimal losses of participants during the trial.
- The doses/exposures of the several contrast groups are adequate to elicit the hypothesized contrast of effects (Figs. 7, 8).
- Intakes of other nutrients or foods, on which the tested nutrient is dependent, must be optimized.

Lacking any of these, an RCT, whether positive or null, will often provide less certainty than well-constructed, prospective observational studies.

In addition, null RCT data should not be construed as indicating that a given food component does not produce a claimed effect unless it can be shown that the dose/exposure was adequate and that the baseline status was low enough to have resulted in the untoward effects that a higher dose would presumably have prevented (Figs. 7, 8).

In case RCT data (or the possibility of developing such) are lacking, alternative lines of evidence should be considered.

- If a proposed claim is supported by two or more well-designed prospective cohort studies, each with an appropriate sample size, if they come from different populations, if all find a similar effect, and if there is a plausible biological basis for the postulated action, such a claim could be approvable.
- Metabolic and nutrigenomic investigations can also provide valuable information regarding mechanisms of action relevant to health outcomes and thus contribute substantially to the biological plausibility of the relation.

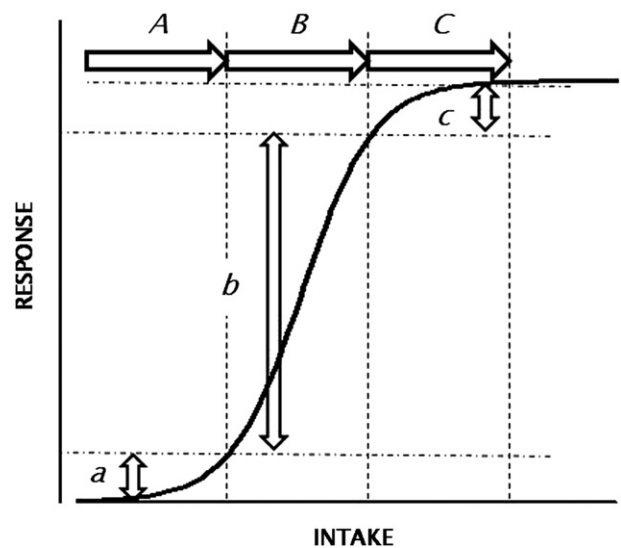


Fig. 7. Typical sigmoid response curve showing the response across different intakes of a given nutrient. Depicted are the expected responses from equal increases in nutrient intake, starting from a low basal intake and moving to progressively higher levels. Intake increments A, B, and C produce responses a, b, and c, respectively. Only intakes in the B region produce large-enough responses to test the hypothesis that the nutrient concerned elicits the response in question. (© 2010 by Robert P. Heaney, M.D. All rights reserved. Used with permission).

- In general, the strength, consistency, and biological plausibility of the evidence should be reflected in the wording of the claim, e.g., claims based on strong evidence or moderate levels of evidence.

2. What are the strengths and limitations of meta-analyses and systematic reviews in nutritional research?

Systematic reviews and meta-analyses are valuable tools summarizing the scientific data available and pointing toward findings that could be regarded as generally accepted. However,

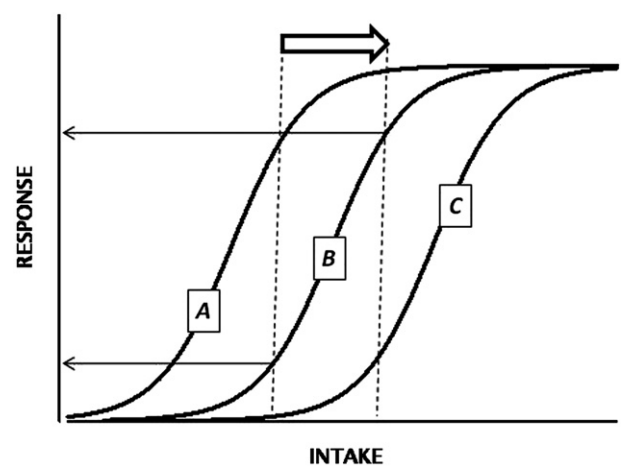


Fig. 8. Three typical sigmoid response curves showing differences in response to any given intake exhibited by three organ systems within the same individual. A given intake increment (open arrow) intersects the three responses in different regions of their respective curves. Thus, the increment shown, starting from a certain basal intake/status, will produce a measurable response in system B, but not in system A or C. For system A, the response occurred at a lower intake, and little more could be expected from further increases. For system C, still higher intakes were needed to produce a perceptible response. (© 2010 by Robert P. Heaney, M.D. All rights reserved. Used with permission).

they need to employ the relevant biological criteria (in addition to the usual methodologic criteria) for the inclusion of studies for estimation of aggregate effects. Reviews and meta-analyses whose outcomes are null and that have included studies that failed to meet the criteria listed earlier should not be accepted as evidence against a claim.

IV. The future of nutritional research

This session addresses whether claims relating to the maintenance of health would need to be assessed in a different way than claims that relate to the decrease of the development of a disease and what is the role of generally accepted scientific evidence in this respect.

Consensus statement

There is a need to address alternative methodologies for the assessment of health effects, where appropriate, targeted to individual physiologic mechanisms.

Over the past century, the research of the relation of diet to health has seen some major transitions, mostly following biochemical and biomedical paths. On the one hand, molecular mechanisms were discovered (energy metabolism, vitamin biochemistry); on the other hand, observations were made on long-term health effects, mostly from a negative (unhealthy, disease) viewpoint. The combination of these two tracks has led to the “paradigm” that optimal nutrition prevents disease and thus disease-related mechanisms and related biomarkers should be used to quantify the effect of a dietary intervention. This mode of working was carried to perfection with the recent boost in biological and biomedical sciences, but in the end essentially demonstrated that this only partly works in nutrition research. Nutrition is not only about preventing disease progression but also about maintaining optimal health and physiologic functions. Nonetheless, this created major difficulties because essentially no “biomarkers of health” were available, and the absence of elevated disease/damage biomarkers was correctly not regarded as indicative for nutritional effects [73].

It is now realized that there needs to be a re-evaluation of the core role of nutrition in relation to physiology. Apart from providing energy, the essential role is to facilitate optimal functioning of all physiologic processes in the context of an ever-changing environment. Physiology is not static but continuously adapting to environmental and chronobiological changes: energy pulses, physical exercise, increased blood pressure, oxidative stress, mental stress, infections, prolonged caloric restriction, changing diurnal rhythms, etc. As an example, a simple glucose diet induces changes not only in glucose and insulin but also in cholesterol, triacylglycerols, oxidative stress, blood pressure, and adiponectin, i.e., all rapidly changing “risk factors” of the metabolic syndrome [74]. The molecular processes underpinning these processes are quantifiable and new “biomarkers” emerge from assays that apply perturbation of homeostasis as a measurement of health [21,75]. With the inclusion of new analytical technologies in these perturbation assays, a wealth of information on the molecular physiology of adaptation emerges, as in the case of the oral glucose tolerance test [76–78].

The application of this concept in various areas of functional adaptation processes will produce new biomarkers of health. Immune function is a typical example, where “homeostatic” values of molecular reporters (cytokines, bioactive lipids) or

functional parameters (cell counts, cellular assays) are non-informative and in fact should not change in a healthy person on a diet that “increases resistance.” However, the recovery time after infection is a clear measurement of immune-related resistance and is used in dietary intervention studies [79].

Perhaps the final step is to skip the word *biomarker*, which has become a paradigm in itself. We are now approaching the stage in biological research where we describe, understand, and quantify the molecular physiologic processes instead of a single parameter in a single point in time that is supposed to capture biological complexity. Nutrition is about health, and once we capture the relevant processes involved in maintaining optimal health, i.e., that continuously adapt under a variety of perturbations, we refocus on the real role of nutrition. In contrast to 100 y ago, we can do this with the help of rapidly increasing biological knowledge and a wide array of newly emerging technologies.

Conclusion

All types and sources of data need to be considered in the weighing of the totality of the evidence, and the best and most logical resource to assess the evidence available in support of the biological plausibility and probability of the beneficial role of a nutrient or food component is by gathering the experts in the field and reaching conclusions in the form of a scientific consensus. This assures that the totality of the evidence is considered and that conclusions are reached in the best knowledge of the current state of scientific research.

The EU PASSCLAIM project supports this approach but provides no indications on how the overall totality of the evidence should be weighed and interpreted. Several different scientific bodies have considered systems for assessing and grading the evidence to establish the strength, consistency, and biological plausibility of a diet–health relation and enable regulators to make decisions on public health recommendations and claims approvals in the best knowledge of the facts. All the various sources of available scientific data should be considered, including observational and mechanistic data, practice of health care professionals, and, in particular for botanicals, evidence of traditional use. Where the evidence is not conclusive, there is a good reason to apply the inverse precautionary principle, meaning that if the evidence is not conclusive but substantially indicative of the effect, then a management decision to allow the claimed effect may be taken for the benefit of the consumer.

The outcome of this consensus conference highlights that the assessment of the relation between intake of a food or food constituent and a beneficial health effect is a complex process. It can be performed based largely on the outcome of RCTs, and it is not disputed that well-conducted RCTs offer the strongest support for cause-and-benefit relations. However, this is the case only if RCTs meet several essential criteria. Furthermore, RCTs are not always possible or available and are seldom capable of addressing the broad field of interactions of the human physiology. RCTs are a difficult tool for nutritional studies, unless the food component of interest has much in common with a pharmaceutical compound. Therefore, the presence or absence of positive RCT data should not be a gate-keeper criterion in the evaluation of a claims assessment process.

Research in the field of nutrition has entered a new stage, where it deserves its own nutritional methodologies that are distinct and more diverse than evidence-based medicine. It requires redefining health in terms of an individual's potential to adapt to internal and

external stimuli rather than using the pharmaceutical paradigm. New ways of addressing nutritional and functional effects of foods and food components are being developed and offer opportunities for future research. Any claims assessment system that is not able to accommodate the wide spectrum of health effects should be adapted to reflect the science available. It should be capable of stimulating these developments toward new research and the development of innovative products and useful information to the benefit of the consumer, including but not limited to public health messages.

Acknowledgments

The authors thank the European Responsible Nutrition Alliance for their financial support of this conference and publishing the report. The authors thank Jana Tinz for her assistance in publishing the report as supplement coordinator. They also thank Dr. Claudia Wicke for the organization of the consensus conference.

References

- [1] Diplock AT, Aggett PJ, Ashwell M, Bornet F, Fern EB, Roberfroid MB. Scientific concepts of functional foods in Europe. Consensus document. *Br J Nutr* 1999;81(suppl 1):S1–27.
- [2] Richardson DP, Affertsholt T, Asp NG, Bruce A, Grossklaus R, Howlett J, et al. PASSCLAIM—synthesis and review of existing processes. *Eur J Nutr* 2003;42(suppl 1):I96–111.
- [3] Hill AB. The environment and disease: association or causation? *Proc R Soc Med* 1965;58:295–300.
- [4] Aggett PJ, Antoine JM, Asp NG, Bellisle F, Contor L, Cummings JH, et al. PASSCLAIM: consensus on criteria. *Eur J Nutr* 2005;44(suppl 1):i5–30.
- [5] Academy of Medical Sciences. Identifying the environmental causes of disease: how should we decide what to believe and when to take action? Report synopsis. London: Academy of Medical Science; 2007. p. 144.
- [6] Blumberg J, Heaney RP, Huncharek M, Scholl T, Stampfer M, Vieth R, et al. Evidence-based criteria in the nutritional context. *Nutr Rev* 2010;68:478–84.
- [7] Howick J, Glasziou P, Aronson JK. The evolution of evidence hierarchies: what can Bradford Hill's 'guidelines for causation' contribute? *J R Soc Med* 2009;102:186–94.
- [8] Howick J, Glasziou P, Aronson JK. Evidence-based mechanistic reasoning. *J R Soc Med* 2010;103:433–41.
- [9] Ward AC. The role of causal criteria in causal inferences: Bradford Hill's "aspects of association." *Epidemiol Perspect Innov* 2009;6:2.
- [10] Phillips CV, Goodman KJ. The missed lessons of Sir Austin Bradford Hill. *Epidemiol Perspect Innov* 2004;1:3.
- [11] Maxwell JA. Using qualitative methods for causal explanation. *Field Methods* 2004;16:243–64.
- [12] World Health Organisation. Diet, nutrition and the prevention of chronic diseases: report of a joint FAG/WHO expert consultation. WHO technical report series 916. Geneva: World Health Organisation; 2004.
- [13] World Cancer Research Fund/American Institute for Cancer Research. Food, nutrition and the prevention of cancer: a global perspective. Washington, DC: World Cancer Research Fund/American Institute for Cancer Research; 1997.
- [14] The Netherlands Ministry of Health. Netherlands proposal for a systematic approach for a generic list of health claims. VWA (Dutch Food Authority) Nutrition Centre. 2004.
- [15] EMEA Working Party on Herbal Medicinal Products. Final concept paper on the implementation of different levels of scientific evidence in core-data for herbal drugs. EMEA/CPMP/HMPWP/1156/03. London; March 3, 2004.
- [16] Höfler M. The Bradford Hill considerations on causality: a counterfactual perspective. *Emerg Themes Epidemiol* 2005;2:11.
- [17] World Health Organization. Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference. 1946.
- [18] Gallagher A, Meijer G, Richardson D, Rondeau V, Skarp M, Stasse-Wolthuis M, et al. A standardised approach towards PROVing the efficacy of foods and food constituents for health CLAIMs (PROCLAIM): providing guidance. *Br J Nutr*; 2011. (in press).
- [19] What is health? The ability to adapt. *Lancet* 2009;373:781.
- [20] van Ommen B, Fairweather-Tait S, Freidig A, Kardinaal A, Scalbert A, Wopereis S. A network biology model of micronutrient related health. *Br J Nutr* 2008;99:S72–80.
- [21] van Ommen B, Keijer J, Heil SG, Kaput J. Challenging homeostasis to define biomarkers for nutrition related health. *Mol Nutr Food Res* 2009;53:795–804.
- [22] Congdon NG, Friedman DS, Lietman T. Important causes of visual impairment in the world today. *JAMA* 2003;290:2057–60.
- [23] Bunce C, Wormald R. Causes of blind certifications in England and Wales: April 1999–March 2000. *Eye (Lond)* 2008;22:905–11.
- [24] Bandello F, Lafuma A, Berdeaux G. Public health impact of neovascular age-related macular degeneration treatments extrapolated from visual acuity. *Invest Ophthalmol Vis Sci* 2007;48:96–103.
- [25] Tomany SC, Wang HJ, van Leeuwen R, Klein R, Mitchell P, Vingerling JR, et al. Risk factors for incident age-related macular degeneration—pooled findings from 3 continents. *Ophthalmology* 2004;111:1280–7.
- [26] AREDS (Age-Related Eye Disease Study) Research Group. A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C and E, beta carotene, and zinc for age-related macular degeneration and vision loss: AREDS report no. 8. *Arch Ophthalmol* 2001;119:1417–36.
- [27] Albanes D, Heinonen OP, Taylor PR, Virtamo J, Edwards BK, Rautalahti M, et al. Alpha-tocopherol and beta-carotene supplements and lung cancer incidence in the alpha-tocopherol, beta-carotene cancer prevention study: effects of base-line characteristics and study compliance. *J Natl Cancer Inst* 1996;88:1560–70.
- [28] Omenn GS, Goodman GE, Thornquist MD, Balmes J, Cullen MR, Glass A, et al. Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease. *N Engl J Med* 1996;334:1150–5.
- [29] Loane E, Kelliher C, Beatty S, Nolan JM. The rationale and evidence base for a protective role of macular pigment in age-related maculopathy. *Br J Ophthalmol* 2008;92:1163–8.
- [30] Bone RA, Landrum JT, Hime GW, Cains A, Zamor J. Stereochemistry of the human macular carotenoids. *Invest Ophthalmol Vis Sci* 1993;34:2033–40.
- [31] Neuringer M, Johnson EJ, Snodderly DM, Sandstrom MM, Schalch W. Supplementation of carotenoid-depleted rhesus monkeys with lutein or zeaxanthin: effects on serum and adipose tissue carotenoids and macular pigment. *Invest Ophthalmol Vis Sci* 2001;42:1209.
- [32] Li B, Ahmed F, Bernstein PS. Studies on the singlet oxygen scavenging mechanism of human macular pigment. *Arch Biochem Biophys* 2010;504:56–60.
- [33] EU Food Information Council. Vitamins: what they do and where to find them 2008. Available at: <http://www.eufic.org/article/en/expid/miniguide-vitamins/>. Accessed June 10, 2011.
- [34] Richer S, Stiles W, Statkute L, Pulido J, Frankowski J, Rudy D, et al. Double-masked, placebo-controlled, randomized trial of lutein and antioxidant supplementation in the intervention of atrophic age-related macular degeneration: the Veterans LAST study (Lutein Antioxidant Supplementation Trial). *Optometry* 2004;75:216–30.
- [35] Jenkins DJ, Wolever TM, Taylor RH, Barker HM, Fielden H, Baldwin JM, et al. Glycemic index of foods: a physiological basis for carbohydrate exchange. *Am J Clin Nutr* 1981;34:362–6.
- [36] Brand-Miller J, Hayne S, Petocz P, Colagiuri S. Low-glycemic index diets in the management of diabetes: a meta-analysis of randomised controlled trials. *Diabetes Care* 2003;26:2261–7.
- [37] Brand-Miller JC, Stockmann K, Atkinson F, Petocz P, Denyer G. Glycemic index, postprandial glycemia, and the shape of the curve in healthy subjects: analysis of a database of more than 1,000 foods. *Am J Clin Nutr* 2009;89:97–105.
- [38] Bjørnholt JV, Erikssen G, Aaser E, Sandvik L, Nitter-Hauge S, Jervell J, et al. Fasting blood glucose: an underestimated risk factor for cardiovascular death. *Diabetes Care* 1999;22:45–9.
- [39] Asia Pacific Cohort Studies Collaboration. Blood glucose and risk of cardiovascular disease in the Asia Pacific region. *Diabetes Care* 2004;27:2836–42.
- [40] Albers R, Antoine JM, Bourdet-Sicard R, Calder PC, Gleeson M, Lesourd B, et al. Markers to measure immunomodulation in human nutrition intervention studies. *Br J Nutr* 2005;94:452–81.
- [41] Berth M, Delanghe J, Langlois M, De Buyzere M. Reference values of serum IgA subclasses in Caucasian adults by immunonephelometry. *Clin Chem* 1999;45:309–10.
- [42] Yel L. Selective IgA deficiency. *J Clin Immunol* 2010;30:10–6.
- [43] Rijkers GT, Bengmark S, Enck P, Haller D, Herz U, Kalliomaki M, et al. Guidance for substantiating the evidence for beneficial effects of probiotics: current status and recommendations for future research. *J Nutr* 2010;140:671S–6S.
- [44] Wolvers D, Antoine JM, Myllyluoma E, Schrezenmeir J, Szajewska H, Rijkers GT. Guidance for substantiating the evidence for beneficial effects of probiotics: prevention and management of infections by probiotics. *J Nutr* 2010;140:698S–712S.
- [45] de Vrese M, Winkler P, Rautenberg P, Harder T, Noah C, Laue C, et al. Effect of *Lactobacillus gasseri* PA 16/8, *Bifidobacterium longum* SP 07/3, *B. bifidum* MF 20/5 on common cold episodes: a double blind, randomized, controlled trial. *Clin Nutr* 2005;24:481–91.

- [46] Turner RB, Bauer R, Woelkart K, Hulsey TC, Gangemi JD. An evaluation of *Echinacea angustifolia* in experimental rhinovirus infections. *N Engl J Med* 2005;353:341–8.
- [47] Boge T, Rémy M, Vaudaine S, Tanguy J, Bourdet-Sicard R, van der Werf S. A probiotic fermented dairy drink improves antibody response to influenza vaccination in the elderly in two randomised controlled trials. *Vaccine* 2009;27:5677–84.
- [48] Damen JE. Harmonization of requirements for influenza virus vaccines: the first year's experience with the EEC guideline. *Biologicals* 1993;21:179–82.
- [49] Feavers I, Knezevic I, Powell M, Griffiths E; WHO Consultation on Serological Criteria for Evaluation and Licensing of New Pneumococcal Vaccines. Challenges in the evaluation and licensing of new pneumococcal vaccines, 7–8 July 2008, Ottawa, Canada. *Vaccine* 2009;27:3681–8.
- [50] Kennedy E, Meyers L. Dietary Reference Intakes: development and uses for assessment of micronutrient status of women—a global perspective. *Am J Clin Nutr* 2005;81:1194S–7S.
- [51] Calvo MS, Whiting SJ. Overview of the proceedings from Experimental Biology 2004 symposium: vitamin D insufficiency: a significant risk factor in chronic diseases and potential disease-specific biomarkers of vitamin D sufficiency. *J Nutr* 2005;135:301–3.
- [52] Institute of Medicine (IOM). Dietary reference intakes for calcium and vitamin D. Washington, DC: National Academic Press; 2011. Available at: <http://www.iolsi.org/Europe/Publications/Beyond%20PASSCLAIM%20Report%202010.pdf>. Accessed June 10, 2011.
- [53] Dalle-Donne I, Rossi R, Colombo R, Giustarini D, Milzani A. Biomarkers of oxidative damage in human disease. *Clin Chem* 2006;52:601–23.
- [54] Fenech M, Baghurst P, Luderer W, Turner J, Record S, Ceppi M, Bonassi S. Low intake of calcium, folate, nicotinic acid, vitamin E, retinol, beta-carotene and high intake of pantothenic acid, biotin and riboflavin are significantly associated with increased genome instability—results from a dietary intake and micronucleus index survey in South Australia. *Carcinogenesis* 2005;26:991–9.
- [55] Blount BC, Mack MM, Wehr CM, MacGregor JT, Hiatt RA, Wang G, et al. Folate deficiency causes uracil misincorporation into human DNA and chromosome breakage: implications for cancer and neuronal damage. *Proc Natl Acad Sci U S A* 1997;94:3290.
- [56] Everson RB, Wehr CM, Ersson GL, MacGregor JT. Association of marginal folate depletion with increased human chromosomal damage in vivo: demonstration by analysis of micronucleated erythrocytes. *J Natl Cancer Inst* 1988;80:525–9.
- [57] Heddle JA, Fenech M, Hayashi M, MacGregor JT. Reflections on the development of micronucleus assays. *Mutagenesis* 2011;26:3–10.
- [58] Schneider M, Diemer K, Engelhart K, Zankl H, Trommer WE, Biesalski HK. Protective effects of vitamins C and E on the number of micronuclei in lymphocytes in smokers and their role in ascorbate free radical formation in plasma. *Free Radic Res* 2001;34:209–19.
- [59] Hartwig A. Role of magnesium in genomic stability. *Mutat Res* 2001;475:113–21.
- [60] AFSSA (Agence Française de Sécurité Sanitaire des Aliments). Framework for the evaluation of the safety, the effect and the claims of foodstuffs made from plants for the human diet 2003. Available at: <http://www.afssa.fr/Documents/NUT-Ra-Plantes.pdf>. Accessed June 10, 2011.
- [61] European Union. Directive 2004/27/EC of the European Parliament and of the Council of 31 March 2004 amending Directive 2001/83/EC on the community code relating to medicinal products for human use. *Off J Eur Union* 2004. L136/34.
- [62] European Union. Corrigendum to Regulation (EC) No 1924/2006 of the European Parliament and of the Council of 20 December 2006 on nutrition and health claims made on foods. *Off J Eur Union* 2006. L12/3.
- [63] International Life Science Institute. Beyond passclaim—guidance to substantiate health claims on foods. ILSI Europe report series. Brussels: International Life Science Institute; 2010.
- [64] Polya G. Biochemical targets of plant bioactive compounds: a pharmacological reference guide to sites of action and biological. New York: CRC Press; 2003.
- [65] Council of Europe. Homeostasis, a practical tool to distinguish between food supplements and medicines 2006. Available at: [http://www.coe.int/t/e/social_cohesion/soc-sp/Homeostasis%20\(2\).pdf](http://www.coe.int/t/e/social_cohesion/soc-sp/Homeostasis%20(2).pdf). Accessed June 10, 2011.
- [66] Pardo-de-Santayana M, Pieroni A, Puri RK, editors. Ethnobotany in the new Europe: people, health and wild plant resources. Oxford: Berghahn Books; 2010.
- [67] Marles R. Aboriginal plant use in Canada's northwest boreal forest. Toronto: UTP; 1999.
- [68] Bao Y, Fenwick R. Phytochemicals in health and disease, Institute of Food Research. Norwich, UK: Marcel Dekker; 2004.
- [69] Surh YJ. Dietary modulation of cell signaling pathways. Boca Raton, London, New York: CRC Press (Taylor & Francis Group); 2009.
- [70] World Health Organisation. WHO monographs on selected medicinal plants. Volumes 1–4. Geneva: World Health Organisation; 1999–2009.
- [71] Teuscher E, Anton R, Lobstein A. Plantes aromatiques: épices, aromates, condiments et leurs huiles essentielles. 1st ed. Paris: Éditions Tec & Doc Lavoisier, Paris; 2005.
- [72] Frohne D, Pfänder HJ, Anton R. Plantes à risques. Un ouvrage destiné aux pharmaciens, médecins, toxicologues et biologistes. 1st ed. Paris: Éditions Tec & Doc Lavoisier; 2009.
- [73] Elliott R, Pico C, Dommels Y, Wybranska I, Hesketh J, Keijer J. Nutrigenomic approaches for benefit-risk analysis of foods and food components: defining markers of health. *Br J Nutr* 2007;98:1095–100.
- [74] Nakatsuji H, Kishida K, Kitamura T, Nakajima C, Funahashi T, Shimomura I. Dysregulation of glucose, insulin, triglyceride, blood pressure, and oxidative stress after an oral glucose tolerance test in men with abdominal obesity. *Metabolism* 2010;59:520–6.
- [75] van Ommen B, Keijer J, Kleemann R, Elliott R, Drevon CA, McArdle H, et al. The challenges for molecular nutrition research 2: quantification of the nutritional phenotype. *Genes Nutr* 2008;3:51–9.
- [76] Shaham O, Wei R, Wang TJ, Ricciardi C, Lewis GD, Vasan RS, et al. Metabolic profiling of the human response to a glucose challenge reveals distinct axes of insulin sensitivity. *Mol Syst Biol* 2008;4:214.
- [77] Spiegel P, Danielsson APH, Bacos K, Nagorny CLF, Moritz T, Mulder H, Filipsson K. Metabolomic analysis of a human oral glucose tolerance test reveals fatty acids as reliable indicators of regulated metabolism. *Metabolomics* 2009;6:56–66.
- [78] Zhao X, Peter A, Fritsche J, Elcnrova M, Fritsche A, Häring HU, et al. Changes of the plasma metabolome during an oral glucose tolerance test: is there more than glucose to look at? *Am J Physiol Endocrinol Metab* 2009;296:E384–93.
- [79] Albers R, van der Wielen RP, Brink EJ, Hendriks HF, Dorovska-Taran VN, Mohede IC. Effects of cis-9, trans-11 and trans-10, cis-12 conjugated linoleic acid (CLA) isomers on immune function in healthy men. *Eur J Clin Nutr* 2003;57:595–603.

Appendix. Various systems proposed for grading of evidence for different types of claims

Function claims [14]

This methodology was established by the Netherlands Ministry of Health and proposed in December 2004 as a systematic approach for a generic list of health claims under the NHCR (Article 13).

Convincing level of evidence would constitute support from
Category 5 evidence: text handbooks, monographs, judgment by government-related organizations, scientific groups, or expert organizations (e.g., WHO, SCF, etc.), critical reviews by independent experts

Category 4 evidence: meta-analysis, peer-reviewed publications, critical reviews by experts

Probable level of evidence would constitute support from
Category 3 evidence: 1 publication of meta-analysis on peer-reviewed articles demonstrating that “positive outweighs the negative,” 1 large human study with supportive epidemiologic and/or laboratory data, multiple small human studies with good designs and consistent results

Insufficient level of evidence would constitute support only from
Category 2 evidence: epidemiologic data, even with consistent results (unless supported by biological plausibility and consistent laboratory data), 1 small human study with supportive laboratory data, 1 large human study with contradictory epidemiologic and/or laboratory data, multiple small human studies showing consistent results but with flawed design or with good designs but showing contradictory results

Category 1 evidence: in vitro or animal (laboratory) data only, 1 small human study or several small uncontrolled human studies, epidemiologic data with contradictory results

Decrease of disease risk claims (World Cancer Research Fund as modified by the WHO [12])

This methodology was used to grade the strength of evidence linking dietary and lifestyle factors to the risk of developing diseases including obesity, type 2 diabetes, cardiovascular disease, cancer, dental diseases, and osteoporosis.

Convincing level of evidence would consist of evidence based on epidemiologic studies showing consistent associations between exposure and disease, with little or no evidence to the contrary. The available evidence is based on a substantial number of studies including prospective observational studies and, where relevant, randomized controlled trials of sufficient size, duration, and quality showing consistent effects. The association should be biologically plausible.

Probable level of evidence: evidence based on epidemiologic studies showing fairly consistent associations between exposure and disease, but where there are perceived shortcomings in the available evidence or some evidence to the contrary, which precludes a more definite judgment

Shortcomings in the evidence may be any of the following: insufficient duration of trials (or studies); insufficient trials (or studies) available; inadequate sample sizes; incomplete follow-up. Laboratory evidence is usually supportive. The association should be biologically plausible.

Possible level of evidence: evidence based mainly on findings from case-control and cross-sectional studies; insufficient randomized controlled trials, observational studies, or non-randomized controlled trials are available; evidence based on non-epidemiologic studies, such as clinical and laboratory investigations, is supportive; more trials are required to support the tentative associations, which should also be biologically plausible

Insufficient level of evidence: evidence based on findings of a few studies that are suggestive but insufficient to establish an association between exposure and disease; limited or no evidence is available from randomized controlled trials; more well-designed research is required to support the tentative associations

Medicinal products*

This methodology for the grading of evidence was developed by the EMEA as a guide for national authorities to grade the level of evidence of published literature submitted in well-established use applications and to assure that the level of evidence and the grading of recommendations correspond to the nature of the disease that is to be treated.

It does not apply to new medicinal products that need specific preclinical tests and clinical trials and to traditional herbal medicinal products that need only demonstration of traditional use if the effect is plausible on the basis of traditional use.

Grade A evidence: requires ≥ 1 randomized controlled trial as part of the body of literature of overall good design and consistency addressing the specific recommendations; it would be expected to come from

Evidence obtained from meta-analysis of randomized controlled trials (category 1a evidence)

Evidence obtained from ≥ 1 randomized controlled trial (category 1b evidence)

Grade B evidence: requires availability of well-conducted clinical studies but no randomized clinical trials on the topic of recommendation. It would be expected to come from

Evidence obtained from ≥ 1 well-designed controlled study without randomization (category 2a evidence)

Evidence obtained from ≥ 1 other type of well-designed quasi-experimental study (category 2b evidence)

Evidence obtained from well-designed non-experimental descriptive studies, such as comparative studies, correlation studies, and case-control studies (category 3 evidence)

Grade C evidence: requires evidence from expert committee reports or opinions and/or clinical experience of respected authorities; indicates absence of directly applicable studies of good quality; it would be expected to come from Evidence obtained from expert committee reports or opinions and/or clinical experience of respected authorities (category 4 evidence)

In addition, the EMEA distinguished different levels of claims for medicinal effects (herbal drugs) [15]:

Major claims (minimum requirements for high level of evidence: level Ia, Ib), e.g., for the treatment, cure, management, or prevention of any serious disease or disorder

Medium claims (minimum requirements for medium level of evidence: level IIa, IIb, III), e.g., decrease of the risk of a disease/disorder; decrease in the frequency of a discrete event; aids/assists in the management of a named symptom/disease/disorder; relief of symptoms of a named disease or disorder

Minor claims (minimum requirements for general level of evidence: level IV), e.g., relief or management of symptoms of a minor, self-limiting disease/disorder that does not require medical intervention for diagnosis or monitoring; description of a pharmacologic action related to management of symptoms of a self-limiting disease/disorder that does not require medical intervention for diagnosis or monitoring

EMEA, European Medicines Agency; NHCR, Nutrition and Health Claims Regulation; SCF, Scientific Committee on Food; WHO, World Health Organization

* EMEA Working Party on Herbal Medicinal Products. Updated draft points to consider on the evidence of safety and efficacy required for well-established herbal medicinal products in bibliographic applications. EMEA/HMPWP/23/99. London, October 25, 1999.